

Antonio Luis Montealegre Gracia

Aplicaciones forestales de los
datos LiDAR-PNOA en ambiente
mediterráneo: su filtrado e
interpolación y el modelado de
parámetros estructurales con
apoyo en trabajo de campo

Departamento
Geografía y Ordenación del Territorio

Director/es
Lamelas Gracia, María Teresa
Riva Fernández, Juan

<http://zaguan.unizar.es/collection/Tesis>



Reconocimiento – NoComercial – SinObraDerivada (by-nc-nd): No se permite un uso comercial de la obra original ni la generación de obras

© Universidad de Zaragoza
Servicio de Publicaciones



Universidad
Zaragoza

Tesis Doctoral

APLICACIONES FORESTALES DE LOS DATOS LIDAR-PNOA EN AMBIENTE MEDITERRÁNEO: SU FILTRADO E INTERPOLACIÓN Y EL MODELADO DE PARÁMETROS ESTRUCTURALES CON APOYO EN TRABAJO DE CAMPO

Autor

Antonio Luis Montealegre Gracia

Director/es

Lamelas Gracia, María Teresa

Riva Fernández, Juan

UNIVERSIDAD DE ZARAGOZA

Geografía y Ordenación del Territorio

2017



TESIS
DOCTORAL

APLICACIONES FORESTALES DE LOS DATOS LIDAR-PNOA EN AMBIENTE MEDITERRÁNEO: SU FILTRADO E INTERPOLACIÓN Y EL MODELADO DE PARÁMETROS ESTRUCTURALES CON APOYO EN TRABAJO DE CAMPO

Antonio Luis Montealegre Gracia

Directores: M^a Teresa Lamelas Gracia y Juan de la Riva Fernández

Zaragoza, 2017



Departamento de
Geografía y
Ordenación del Territorio
Universidad Zaragoza



El autor de esta tesis doctoral contó para su realización con la financiación del “Programa de ayudas para la formación y contratación de personal investigador del Gobierno de Aragón”, Departamento de Ciencia, Tecnología y Universidad (referencia FPI BOA 30, 11/02/2011), así como del Proyecto de Investigación del Centro Universitario de la Defensa de Zaragoza (Proyecto N° 2013-04). Dejo aquí constancia de mi agradecimiento.

De acuerdo con la modalidad ofrecida por el Programa de Doctorado en Ordenación del Territorio y Medio Ambiente de la Universidad de Zaragoza, se presenta esta tesis doctoral como un compendio de publicaciones. El doctorando, Antonio Luis Montealegre Gracia, figura como primer autor y responsable de todos y cada uno de los artículos publicados. A continuación se detallan las referencias completas de los artículos que constituyen el cuerpo de la tesis:

1. **Montealegre, A.L.**, Lamelas, M.T., de la Riva, J. 2015. Comparison of open source LiDAR filtering algorithms in a Mediterranean forest environment. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 8: 4072-4085. doi:10.1109/JSTARS.2015.2436974.
2. **Montealegre, A.L.**, Lamelas, M.T., de la Riva, J. 2015. Interpolation routines assessment in ALS-derived Digital Elevation Models for forestry applications. *Remote Sensing*, 7: 8631-8654. doi:10.3390/rs70708631.
3. **Montealegre, A.L.**, Lamelas, M.T., de la Riva, J., García-Martín, A., Escribano, F. 2016. Use of low point density ALS data to estimate stand-level structural variables in Mediterranean Aleppo pine forest. *Forestry*, 89 (4): 373-382. doi:10.1093/forestry/cpw008.
4. **Montealegre, A.L.**, Lamelas, M.T. Tanase, M., de la Riva, J. 2014. Forest fire severity assessment using ALS data in a Mediterranean environment. *Remote Sensing*, 6: 4240-4265. doi:10.3390/rs6054240.

A mis abuelos

Agradecimientos

En primer lugar, quiero expresar mi más sincero agradecimiento a mis directores de tesis, la Dra. María Teresa Lamelas Gracia y el Dr. Juan de la Riva Fernández, no sólo por su ayuda y dedicación constante en la realización de este trabajo de investigación, sino también por todos sus consejos y sabias contribuciones que me han hecho crecer profesional y personalmente a lo largo de estos últimos años. Sin ellos jamás hubiera podido llevar a cabo todo este trabajo.

Asimismo, quiero dar las gracias a los coautores de las publicaciones que forman parte de esta tesis: el Dr. Mihai Tanase, el Dr. Alberto García-Martín y el Tte. Col. Francisco Escribano Bernal, sobre todo por sus enriquecedoras aportaciones, por su saber hacer y por el esfuerzo dedicado en las intensísimas jornadas de trabajo de campo.

Quisiera extender mis agradecimientos al Grupo de Investigación GEOFOREST-IUCA en particular, y en general al Departamento de Geografía y Ordenación del Territorio y al Centro Universitario de la Defensa de Zaragoza por darme la oportunidad de poder llevar a cabo mi proceso formativo como doctorando, por apoyarme con sus recursos materiales y económicos, y en definitiva, por ayudarme a desarrollar esta investigación. Me siento un verdadero privilegiado por haber formado parte de una gran familia compuesta por personas de una gran calidad humana.

Agradezco también las palabras de aliento de mis compañeros, doctorandos y becarios que me han hecho más llevadera la fatiga en esta carrera de fondo. Dhais Peña, Daniel Borini, Olga Rosero, Lidia Vlassova, Darío Domingo y Adrián Jiménez, mil gracias por todo.

Del mismo modo, doy gracias al caprichoso destino que, en aquel congreso internacional de Teledetección ambiental en Berlín, hizo que se cruzara en mi camino la Dra. Katalin Varga, con la cual he compartido, al lado de Daniel, Olga y Darío, muy buenas experiencias y momentos.

Me gustaría dar las gracias a dos excelentes personas con las cuales he tenido el placer de trabajar durante este último año de doctorado: la Dra. María Sebastián López y el doctorando Carlos López Escolano. Para mí habéis sido realmente el *"Dream Team"*.

Quisiera tener un recuerdo de gratitud hacia mis profesores y mentores del Departamento de Geografía y Ordenación del Territorio que me formaron como geógrafo, y agradecer a todos mis alumnos que, sin ser conscientes de ello, me han enseñado mucho sobre la labor docente, una de las actividades que más me puede satisfacer en la vida.

Finalmente, mi más profundo agradecimiento a mi familia, particularmente a mi madre y a mi padre por todo lo que significan para mí, por su amor, por su apoyo, por sus sacrificios, por entenderme y por permitirme cumplir este sueño.

GRACIAS a todos de corazón.

Antes de que surgiera la tecnología LiDAR (*Light Detection and Ranging*) aeroportada comercial de pulsos discretos y huella pequeña –también conocida como ALS (*Airborne Laser Scanner*)– allá por los años 90 del siglo XX y el uso de los datos fuera accesible para los investigadores en materia forestal, la mayoría de los estudios para derivar parámetros estructurales de la vegetación utilizaban sensores de teledetección ópticos –que solo proporcionan información espectral sobre la cubierta vegetal– y en menor medida sensores SAR (*Synthetic Aperture Radar*) –que aportan información relacionada con la estructura del dosel– debido a la sensibilidad de las microondas. Sin embargo, con la disponibilidad de datos LiDAR capturados para todo el territorio español en el marco del Plan Nacional de Ortofotografía Área (PNOA) –desde el año 2011 para la Comunidad Autónoma de Aragón– se ha abierto una nueva y fructífera línea de investigación.

En esta tesis doctoral se ha hecho uso de dichos datos públicos para estimar parámetros estructurales relacionados con la vegetación, con apoyo en trabajo de campo para la calibración y la validación de los resultados. El tema de trabajo incorpora, además, dos aspectos que le confieren dificultad y complejidad, pero que, al mismo tiempo, le dotan de un interés añadido: en relación con los datos utilizados, la baja densidad de las nubes de puntos; respecto del área de estudio, sus características ambientales típicamente mediterráneas, de topografía accidentada y con formaciones forestales perennifolias, morfológicamente heterogéneas.

La tesis se presenta como un compendio de cuatro artículos científicos que secuencialmente dan respuesta al objetivo principal planteado; a ellos se han añadido dos comunicaciones a congresos que profundizan todavía más en él. En primer lugar, se realizó un análisis comparativo de algoritmos de filtrado implementados en *software* libre con objeto de seleccionar el más adecuado para el área de estudio. En segundo lugar, se investigó sobre la idoneidad de los métodos de interpolación para generar modelos digitales de elevaciones (MDE), necesarios para la normalización de la altura de la nube de puntos. En tercer lugar, se modelaron parámetros estructurales de la vegetación a partir de métodos de regresión lineal y logística múltiples. Estas técnicas permiten, por una parte, el ajuste de modelos basados en la relación de datos estimados en parcelas de campo y datos derivados de las nubes de puntos LiDAR y, por otra, la cartografía de un conjunto de variables dasométricas (altura media de la masa forestal, diámetro cuadrático medio de los fustes, área basimétrica, volumen, densidad de pies de arbolado, ratio de copa y biomasa) y el impacto del fuego forestal sobre la estructura de la vegetación (severidad post-fuego). Por último se generó, mediante técnicas de clasificación digital, un mapa de alta resolución espacial de modelos de combustible forestal basado en la combinación de datos LiDAR y de una imagen óptica multiespectral.

Los resultados y conclusiones que se derivan de la investigación indican que la utilización de filtros basados en superficies, especialmente del algoritmo de curvatura multiescala, así como de métodos de interpolación como el “punto-TIN-raster” constituyen la mejor opción, frente al resto de metodologías existentes, para tratar nubes de puntos de baja densidad en ambientes mediterráneos, si bien parece que existe una cierta dependencia en relación con las características particulares de la zona de estudio. La aplicación de modelos de regresión lineal y logística ha revelado la existencia de una elevada relación entre el dato de campo y la capacidad explicativa de las variables derivadas de los retornos LiDAR relacionadas con la estructura de la vegetación, lo que ha permitido estimar diversas variables

de inventario forestal a nivel de masa, útiles para los gestores del bosque, así como cartografiar el impacto del fuego sobre la estructura de la vegetación. Por otra parte, a la espera de la madurez en el desarrollo de nuevos sensores LiDAR multiespectrales, la combinación de sensores activos y pasivos es una buena alternativa para la obtención de una cartografía precisa sobre los tipos de combustibles forestales de un ecosistema mediterráneo continental.

Previous to the development, back in the 90's of the 20th century, of the commercial technology based on small-footprint discrete pulses LiDAR, also known as Airborne Laser Scanner (ALS), and the availability of data by researchers in forestry, most studies to derive structural parameters of vegetation used optical remote sensing sensors -which only provide spectral information on the vegetation cover-, and to a lesser extent SAR (Synthetic Aperture Radar) sensors -which provide information related to the structure of the canopy- due to the sensitivity of microwaves. However, the availability of LiDAR data for the whole Spain territory within the framework of the National Plan for Aerial Orthography (PNOA) -since 2011 for the Autonomous Community of Aragón- has opened a new and fruitful research activity.

In this doctoral thesis has been made use of that public data to estimate vegetation structural parameters, with support in field work for the calibration and validation of the results. This research topic is of relevance for two fundamental questions: the low density of point clouds and the typically Mediterranean environmental characteristics of the study area, which is defined by a rugged topography and evergreen and morphologically heterogeneous forest formations.

The thesis is presented as a compendium of four scientific articles that sequentially answer to the main objective posed, along with two proceedings that go even deeper into it. Firstly, a comparative analysis of filtering algorithms implemented in free software was performed in order to select the most suitable for the study area. Second, we investigated the suitability of interpolation methods to generate digital elevation models (MDE), which are necessary for the normalization of the point cloud height. Third, vegetation structural parameters were modeled using multiple linear and logistic regression methods. These techniques allow, on the one hand, the adjustment of models based on the relationship of estimated data in field plots and data derived from the LiDAR point cloud and, on the other hand, the mapping of a set of dasometric variables (mean stand height, stem diameter, basal area, volume, stand density, crown ratio, and biomass), and the impact of forest fire on the vegetation structure (post-fire severity). Finally, a high resolution map of forest fuel models based on the combination of LiDAR data and a multispectral optical image was generated using digital classification techniques.

The results and conclusions derived from the research indicate that the use of surface-based filters, especially the multiscale curvature algorithm, as well as interpolation methods such as "point-TIN-raster", are the best option to manage low density point clouds in Mediterranean environments. However, it seems to be certain dependence on the particular characteristics of the study area. The application of linear and logistic regression models has revealed the existence of a strong relationship between the field data and the variables derived from the LiDAR returns, attesting the explanatory capacity of these variables related to the vegetation structure and allowing the estimation of several inventory variables, useful for forest managers, as well as mapping the impact of fire on vegetation structure. On the other hand, waiting for the maturity in the development of new multispectral LiDAR sensors, the combination of active and passive remote sensing is a good alternative for accurate mapping of forest fuel types in Mediterranean continental ecosystems.

1	Introducción.....	1
1.1	<i>Estado de la cuestión.....</i>	<i>1</i>
1.2	<i>Justificación.....</i>	<i>16</i>
1.3	<i>Hipótesis y objetivos.....</i>	<i>17</i>
1.4	<i>Estructura.....</i>	<i>18</i>
2	Área de estudio, material y métodos.....	23
2.1	<i>Área de estudio.....</i>	<i>23</i>
2.2	<i>Material y métodos.....</i>	<i>28</i>
2.2.1	<i>Datos LiDAR-PNOA.....</i>	<i>28</i>
2.2.2	<i>Algoritmos de filtrado.....</i>	<i>34</i>
2.2.3	<i>Métodos de interpolación espacial.....</i>	<i>38</i>
2.2.4	<i>Validación de los métodos de filtrado e interpolación y análisis del error.....</i>	<i>41</i>
2.2.5	<i>Inventario forestal y estimación de variables dasométricas.....</i>	<i>46</i>
2.2.6	<i>Estimación de la severidad post-incendio en campo.....</i>	<i>49</i>
2.2.7	<i>Identificación en campo de combustibles forestales.....</i>	<i>50</i>
2.2.8	<i>Imágenes ópticas e índices espectrales.....</i>	<i>52</i>
2.2.9	<i>Variables derivadas de las nubes de puntos LiDAR.....</i>	<i>55</i>
2.2.10	<i>Modelos de regresión.....</i>	<i>58</i>
2.2.11	<i>Representación cartográfica de las variables estimadas.....</i>	<i>63</i>
2.2.12	<i>Clasificación digital basada en la combinación de datos.....</i>	<i>63</i>
3	Aportación de las investigaciones realizadas.....	67
3.1	<i>Comparación de métodos de filtrado para la clasificación de las nubes de puntos LiDAR del PNOA.....</i>	<i>69</i>
3.2	<i>Comparación de métodos de interpolación para la generación de MDE en espacios forestales.....</i>	<i>85</i>
3.3	<i>Estimación de variables dasométricas en masas de <i>P. halepensis</i> Mill.</i>	<i>111</i>
3.4	<i>Estimación del carbono contenido en la biomasa arbórea.....</i>	<i>123</i>
3.5	<i>Modelo logístico para la evaluación de la severidad post-incendio.....</i>	<i>137</i>
3.6	<i>Cartografía de tipos de combustible forestal.....</i>	<i>165</i>
4	Conclusiones y líneas de trabajo futuras.....	175
4.1	<i>Conclusiones principales.....</i>	<i>177</i>
4.2	<i>Conclusiones específicas.....</i>	<i>179</i>
4.3	<i>Líneas de trabajo futuras.....</i>	<i>182</i>
	Referencias.....	185
	Apéndice.....	199

1 INTRODUCCIÓN

Este capítulo describe el marco teórico de la investigación y los conceptos fundamentales que se han tenido en cuenta para desarrollar la tesis doctoral. En primer lugar, se presenta la tecnología láser escáner –singularmente aeroportada y de pulsos discretos–, así como los algoritmos para el filtrado y la interpolación de la nube de puntos orientados a la generación de modelos digitales de elevaciones (MDE) en entornos forestales, teniendo en cuenta las posibles fuentes de error que implican ambos procesos. Posteriormente, se exponen los antecedentes relacionados con la extracción de variables de interés forestal a partir de datos LiDAR en el contexto de las técnicas de teledetección. Todo ello permite conformar la justificación de la investigación y sentar las bases de la hipótesis y los objetivos a alcanzar. Por último, el capítulo finaliza con una explicación sobre la estructura de la tesis doctoral, donde se muestran los trabajos realizados, sus interrelaciones y, en definitiva, la unidad temática del documento.

1.1 Estado de la cuestión

Las técnicas de teledetección son utilizadas actualmente por disciplinas diversas, como la geografía, la geología, la ingeniería civil, las ciencias forestales, la agricultura, la oceanografía, la meteorología, etc. para analizar diversos fenómenos que ocurren en la superficie terrestre. Tradicionalmente se han empleado las imágenes de satélite y las ortofotografías aéreas, sin embargo, en los últimos veinte años la tecnología de escaneo láser ha cobrado gran importancia debido a su capacidad para recoger datos espaciales georreferenciados de forma automática y eficiente (Wehr and Lohr 1999). Esta tecnología de escaneo láser también se conoce con el acrónimo de LiDAR (*Light Detection and Ranging*) (Reutebuch *et al.* 2005), porque se trata de un dispositivo que permite determinar la distancia entre un emisor láser y un objeto o superficie utilizando un haz de luz monocromático, coherente y direccional.

Los antecedentes históricos de la exploración láser se remontan a 1960, con la invención del primer láser de rubí desarrollado por Theodore Harold Maiman que emitía potentes pulsos de luz roja colimada. El progreso de la tecnología “láser”, acrónimo de *light amplification by stimulated emission of radiation*, verdaderamente comenzó en la década siguiente, cuando se perfeccionó la técnica y fueron surgiendo muchas aplicaciones prácticas fuera del laboratorio. Así, los estudios relacionados con los espacios forestales se iniciaron a finales de la década de 1990, con la generación de MDE, la estimación de la altura del dosel arbóreo y del volumen, y la localización e individualización de árboles (Lim *et al.* 2003, Hyyppä *et al.* 2004, Hyyppä *et al.* 2006). Desde los últimos 15 años, la tecnología LiDAR –que se encuentra hoy en una etapa de madurez en cuanto a *hardware*– ha experimentado un crecimiento exponencial debido a sus amplias posibilidades de aplicación en el ámbito forestal, favorecido por su

entrada en el sector comercial y por el desarrollo de *software* específico para el tratamiento de los datos (Renslow 2013).

El escáner láser aerotransportado en aviones o helicópteros, también conocido como ALS (*Airborne Láser Scanning*), es uno de los tipos de LiDAR más comunes (Maltamo *et al.* 2014) en comparación con los sistemas de escaneo sobre plataformas terrestres (trípodes, vehículos) y espaciales (satélites). En el caso concreto de los sistemas aeroportados, la altura de vuelo habitualmente está en el rango de 200 a 300 m para helicópteros y 500 a 1.000 m para aviones (Baltasvias 1999a,c). Aquellos que están orientados al escaneo de la superficie terrestre operan habitualmente en longitudes de onda del infrarrojo próximo (900 a 1.064 nm), donde la reflectividad de la vegetación y la transmisividad de la atmósfera son altas (Lefsky *et al.* 2002). Se trata, por tanto, de un sistema que emite su propio flujo electromagnético, por lo que es independiente de las condiciones de iluminación solar, siempre que la atmósfera esté libre de humo, niebla densa o nubes que puedan interceptar o dispersar la trayectoria del haz láser emitido, ya que la luz infrarroja no penetra en el vapor de agua (Baltasvias 1999a, Vosselman and Maas 2010).

Las nubes de puntos son el producto principal resultante de cualquier sistema LiDAR. Se trata de un conjunto denso de datos que reúne todas las reflexiones (ecos o retornos) producidas por los objetos al ser impactados por el haz de luz láser. A la hora de medir la distancia que separa el emisor de pulsos láser del objeto o superficie se pueden distinguir dos tipologías (Wehr and Lohr 1999, Lim *et al.* 2003, Vosselman and Maas 2010, Wulder *et al.* 2012):

- Sistemas de onda continua (*full-waveform*), que registran toda la energía reflejada (forma de onda). El valor de distancia emisor-objeto se obtiene mediante la medición de la diferencia de fase entre la señal emitida y la radiación recibida por retrodispersión procedente de la superficie del objeto.
- Sistemas de pulsos discretos, que registran uno o varios retornos por cada pulso láser emitido. La distancia recorrida por el fotón de luz desde el trasmisor al objeto se calcula del siguiente modo (ecuación 1):

$$D = v \times \frac{t}{2} \quad (1)$$

Donde v es la velocidad de pulso de láser (que se supone igual a la velocidad de la luz en el vacío, aproximadamente 0,3 m/ns) y t es el tiempo que tarda el pulso láser en regresar al emisor desde que es emitido y reflejado por el objeto (Baltasvias 1999b) (Figura 1).

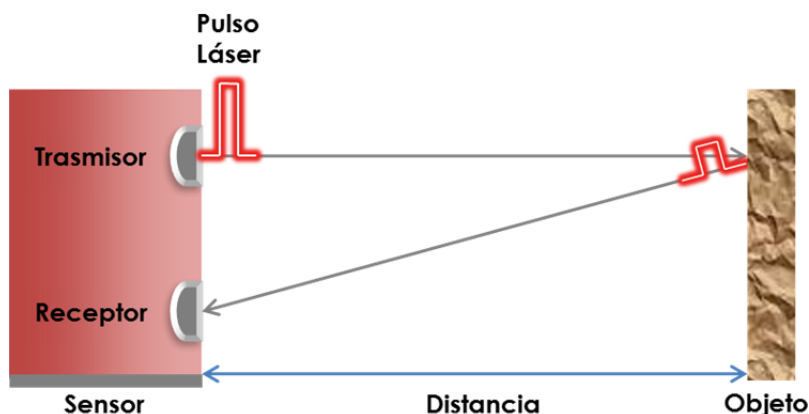


Figura 1. Esquema del funcionamiento de los sistemas de pulsos discretos.

Son precisamente estos últimos sistemas los que han tenido un amplio calado en el sector comercial y los más utilizados en aplicaciones topográficas y forestales (Shan and Toth 2008). No obstante, los sistemas de onda continua, a pesar de la complejidad que implica su uso, están cobrando cada vez más protagonismo debido a la riqueza que ofrecen los datos para caracterizar los objetos. Los primeros sistemas solo eran capaces de grabar un solo retorno o el primero y último por cada pulso láser emitido, pero los sistemas más modernos para aplicaciones topográficas y ambientales o “multi-retorno” pueden registrar hasta cinco retornos por pulso (Lefsky *et al.* 2002). En estos sistemas, cuando el pulso es interceptado por un objeto, una parte de la energía se refleja hacia el receptor y se registra como el primer eco. Pero cuando el objeto es “poroso”, como es el caso de la vegetación, y no bloquea completamente el pulso, la energía continúa su trayectoria, pudiendo ser reflejada no solo por los objetos más altos (primer retorno correspondiente a la superficie del dosel), sino también por los intermedios y bajos (segundo, tercero o cuarto retorno relativos a las ramas, hojas y sotobosque) e incluso por el suelo (quinto y último retorno) (Reutebuch *et al.* 2005). Este hecho es el que se produce cuando el LiDAR aeroportado captura datos de una zona forestal, donde las copas de los árboles presentan pequeños huecos entre las ramas y el follaje por donde los pulsos láser pueden colarse (ver Figura 2).

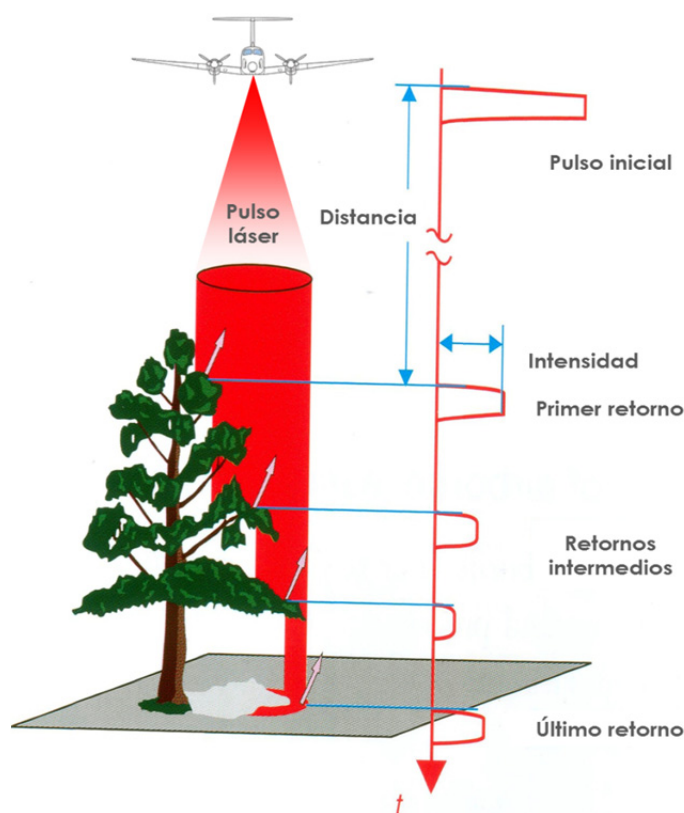


Figura 2. Interacción de un pulso láser con la vegetación. Modificada y traducida de Vosselman y Maas (2010).

Aunque pueda pensarse que el último retorno siempre corresponde a la reflexión producida por el suelo, en la práctica, la densidad de la cubierta vegetal juega un papel decisivo. Según Chasmer *et al.* (2006), sólo el 50% de los últimos retornos en zonas forestales proceden del suelo, por lo que es necesario determinar qué retornos pertenecen realmente a la superficie del suelo utilizando diferentes técnicas de clasificación. Por ejemplo, cuando el objetivo principal es producir un MDE de una zona boscosa, la mayoría de las campañas de vuelo tratan de efectuarse en condiciones de

ausencia de follaje (invierno) –si se trata de especies de hoja caduca– para maximizar el porcentaje de pulsos reflejados por el suelo. Por el contrario, cuando el objetivo principal es la caracterización de la estructura de la vegetación, el vuelo debe realizarse en condiciones de máximo desarrollo foliar para incrementar el número de retornos procedentes de las copas de los árboles y del sotobosque (Reutebuch *et al.* 2005).

En esencia, cada pulso láser tiene un diámetro o “huella” que hace referencia a la superficie reflectante iluminada y que depende de la divergencia del haz y de la altura de vuelo (Shan and Toth 2008, Vosselman and Maas 2010). Generalmente, en los sistemas de pulsos discretos oscila entre 0,2 y 1,0 m, aunque existen sistemas con un diámetro de huella superior y que son considerados de huella grande (Reutebuch *et al.* 2005, Evans *et al.* 2009). Es decir, cada pulso puede asemejarse a un cilindro de luz, ya que además presenta una longitud definida por el lapso de tiempo en el que el emisor se activa y desactiva. Por tanto, todas las superficies reflectantes dentro de la huella del pulso generan una onda de energía de retorno, conocida como intensidad de la señal, que es registrada por el sistema en 8 (valores de 0 a 255) o 12 bits (0 a 4095). Esa información puede representarse en una imagen ráster en escala de grises (Figura 3), simulando una fotografía aérea; sin embargo –debido a los múltiples factores que influyen en la intensidad registrada por el sensor, como la altura de vuelo, las condiciones atmosféricas, la reflectividad y el color del objeto, y los ajustes del instrumento láser– dicha imagen hay que utilizarla con cautela para propósitos de clasificación digital (Baltsavias 1999a,b).

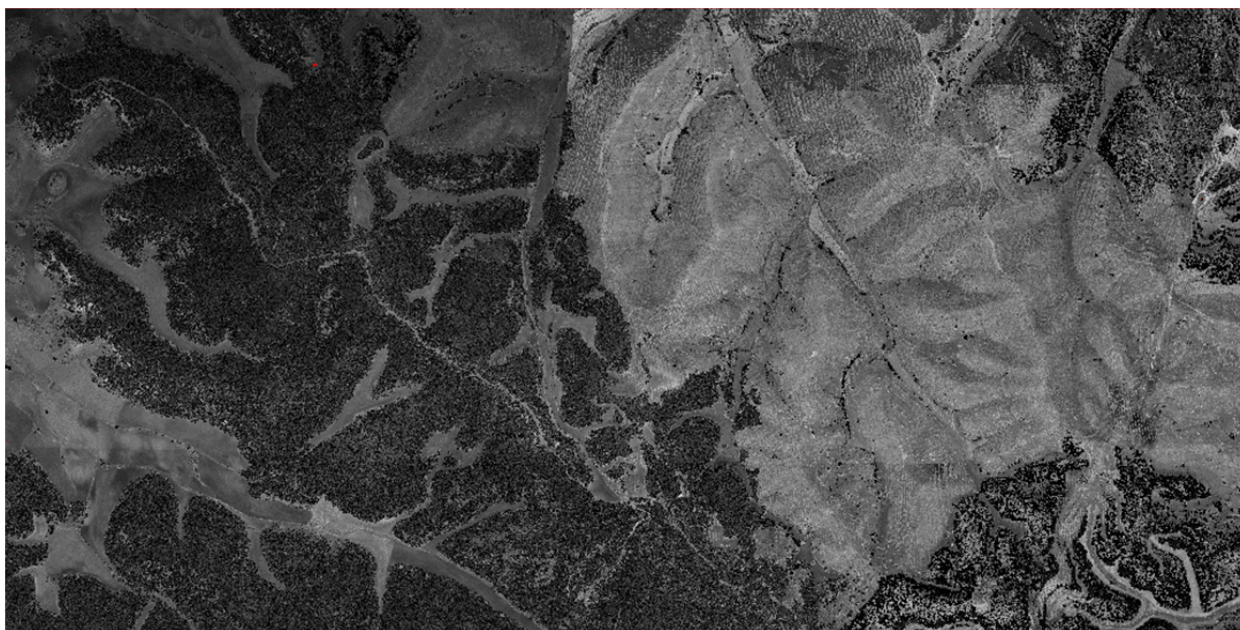


Figura 3. Ejemplo de imagen de intensidad creada a partir de los retornos láser. Se recoge una zona forestal próxima a Zuera (Zaragoza), donde los píxeles más oscuros pertenecen a arbolado y los más claros a zonas de suelo desnudo o vegetación escasa y cultivos en secano o abandonados.

Los principales componentes de un escáner láser de pulsos discretos son: el emisor de pulsos láser, el mecanismo de escaneo y el receptor para registrar y medir la distancia al objeto. Durante el vuelo, el sistema emite pulsos con una duración de unos pocos nanosegundos (10^{-9} s) y una alta tasa de frecuencia de escaneo¹ (hasta 300 kHz), dirigiéndolos hacia la superficie terrestre, la cual se

¹ La frecuencia de escaneo, frecuencia de repetición de pulsos o frecuencia de pulso, hace referencia al número de pulsos por segundo enviados, medida habitualmente en kHz. Por ejemplo, 1 kHz es igual a 1.000 pulsos por segundo.

va escaneando en la dirección dada por el mecanismo de exploración y a medida que avanza la aeronave (Carter *et al.* 2012). Las líneas de escaneado sobre la superficie, perpendiculares a la dirección de vuelo, forman patrones específicos dependiendo del mecanismo de exploración que desvía el haz de luz láser emitido por el escáner, por ejemplo, en forma de dientes de sierra si se trata de un espejo oscilante, líneas paralelas si es un prisma rotatorio, etc. (Vosselman and Maas 2010). La franja de barrido o exploración, es decir, la anchura de la zona que puede ser cubierta durante el vuelo en una dirección, es determinada tanto por la configuración del ángulo de escaneo del sensor o el campo de visión (FOV-Field Of View), como por la altura de la aeronave (Vosselman and Maas 2010). Para cubrir grandes extensiones de superficie, la aeronave realiza varias pasadas paralelas con un cierto solape entre ellas (hasta un 50% de la superficie o más) con la finalidad de que no haya vacíos de información entre ellas (Evans *et al.* 2009).

Cabe señalar que, por sí solos, los retornos no son suficientes para construir un MDE o derivar otro tipo de productos con una referencia espacial explícita. Para solventar este inconveniente, el equipo se completa (ver Figura 4) con un sistema de posicionamiento global (GPS-Global Positioning System) y un sistema inercial de navegación (INS-Inertial Navigation System) con una unidad de medición inercial IMU (Inertial Measurement Unit) (Renslow 2013). De este modo, las coordenadas tridimensionales (x, y, z) de los retornos se calculan a partir de la posición precisa del escáner, determinada por el GPS diferencial con respecto a las estaciones base GPS terrestres, mientras que los ángulos de navegación del avión (cabeceo, alabeo y guiñada) son medidos por la IMU (Baltsavias 1999b, Wehr and Lohr 1999). La combinación, mediante *software* especializado, de la información procedente del escáner láser, GPS e IMU permite asignar coordenadas a cada retorno registrado por el escáner láser con una alta precisión (Li *et al.* 2004), lo que permite conformar la nube de puntos (Aguilar and Mills 2008, Baltsavias 1999b,c, Wehr and Lohr 1999, Carter *et al.* 2012).

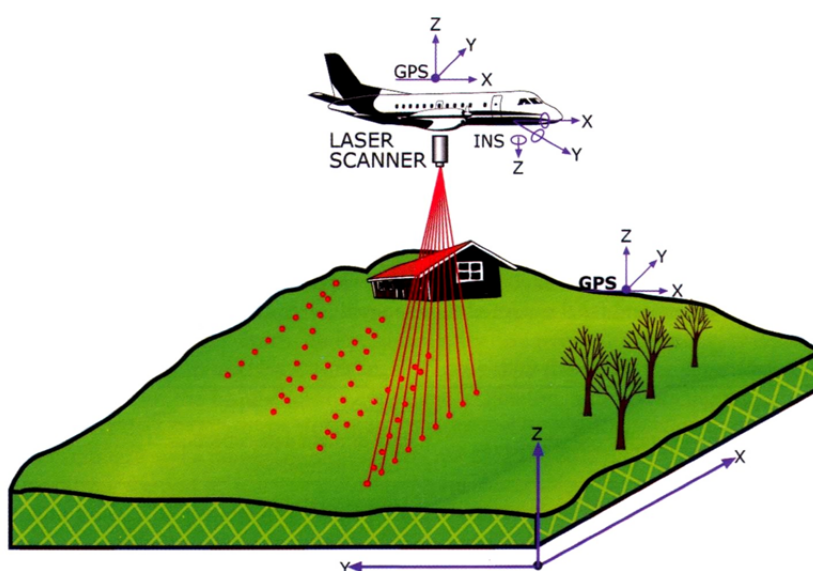


Figura 4. Componentes del LiDAR aeroportado. Tomado de Renslow (2013).

La densidad de puntos de la nube es la manera de expresar la resolución espacial de un conjunto de datos LiDAR en número de puntos por m^2 . La densidad presenta una relación directa con los costes de adquisición y con la calidad y la precisión de los productos resultantes. Habitualmente oscila entre 0,3 y 20 puntos/ m^2 , dependiendo de la velocidad del avión, la altitud y el ángulo de escaneado (Gatzolis and Andersen 2008, Evans *et al.* 2009). De acuerdo con Evans *et al.* (2009), las

densidades de 4-6 puntos/m² proporcionan un buen acuerdo entre el coste y la exactitud en las aplicaciones relacionadas con la vegetación, mientras que algunos proveedores optan por densidades inferiores, en torno a 1 punto/m², para abaratar costes sin comprometer la obtención de resultados aceptables. Por otro lado, la mayoría de los proveedores de sistemas aeroportados proporcionan nubes de puntos con valores de error cuadrático medio entre 5-15 cm en altimetría y 25-100 cm en planimetría (Vosselman and Maas 2010), lo que evidencia una alta precisión en el conjunto de datos.

En comparación con los sistemas tradicionales de topografía y fotogrametría, la tecnología ALS proporciona una alternativa rápida y directa para reunir una gran cantidad de mediciones tridimensionales en formato digital de la superficie terrestre con una alta resolución y precisión, razón por la cual se ha adoptado gradualmente como la técnica principal para generar MDE (Baltsavias 1999a, Hill *et al.* 2000, Polat *et al.* 2015). De acuerdo con Li *et al.* (2004), un MDE permite describir las alturas o elevaciones del terreno referidas a un datum vertical común mediante una representación que refleja el relieve a través de valores numéricos que simbolizan la cota o altura.

El desarrollo de los MDE involucró inicialmente a fotogrametristas e ingenieros civiles, pero el salto importante se produjo cuando los científicos en geometría computacional, matemática aplicada e informática comenzaron a desarrollar algoritmos de modelado y sistemas para gestionar los datos. En último término, los especialistas de diversas disciplinas han jugado un papel clave en el uso de los MDE para diferentes propósitos. Por ejemplo, a partir de la década de 1990, con el desarrollo de sistemas de información geográfica (SIG), el MDE se empezó a utilizar cada vez más en el ámbito de las ciencias ambientales.

Por consiguiente, el modelado digital del terreno comprende cuatro componentes principales de acuerdo con Li *et al.* (2004): la adquisición, el proceso computacional, la gestión de los datos y el desarrollo de aplicaciones (ver Figura 5).

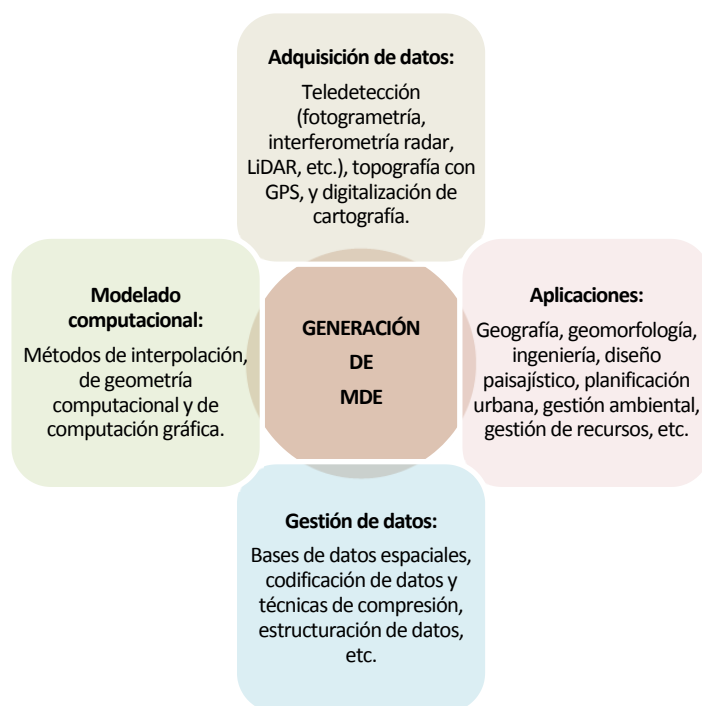


Figura 5. Relaciones entre los componentes que intervienen en el modelado digital del terreno a través de la generación de MDE.

Tal como se muestra en la Tabla 1, existen diferentes técnicas –directas e indirectas– para la captura de la variable elevación; sin embargo, la principal ventaja de los MDE derivados de los sistemas LiDAR es que no se ven afectados por sombras causadas por la vegetación, como les ocurre a los obtenidos mediante fotogrametría (Hyde *et al.* 2006). Aunque la precisión del MDE disminuye en zonas de vegetación densa, es mejor a la obtenida mediante restitución, ya que entre el 20-40% de los pulsos pueden alcanzar el suelo (Ackermann 1999, Hodgson *et al.* 2005). Además, a diferencia de lo que ocurre en fotogrametría, no es necesario que los puntos aparezcan en dos imágenes (Polat and Uysal 2015), evitando la redundancia de información. Por el contrario, los sistemas ALS tienen la desventaja de que en superficies cubiertas por cuerpos de agua –especialmente si son profundos– el haz de energía no retorna al sensor, dando lugar a vacíos de información (Huisman and Gomes Pereira 1998, Vosselman and Maas 2010).

Tabla 1. Principales técnicas de adquisición de datos para la generación de MDE (Baltasavias 1999a, Li *et al.* 2004, Maguya *et al.* 2013).

Técnica/Características	Teledetección	Topografía	Digitalización
Método de adquisición	La fotogrametría y la interferometría radar son métodos indirectos, mientras que el láser mide la elevación directamente	Directo	Indirecto
Precisión de la medición	Centimétrica, dependiendo de la resolución espacial de las imágenes utilizadas	Milimétrica	Métrica
Eficiencia	Alta, dada la automatización de los procesos. Método útil para áreas medianas-grandes	Muy intensivo en términos de mano de obra. Método adecuado para áreas pequeñas y de fácil acceso	Alta, dada la automatización de los procesos. Método útil para áreas grandes

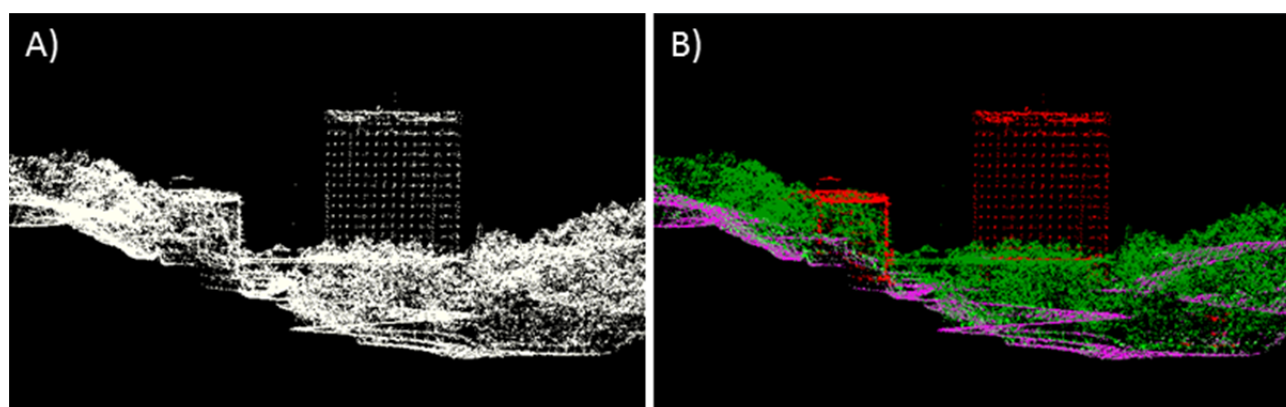


Figura 6. A) Ejemplo de nube de puntos sin filtrar capturada por un sistema ALS correspondiente a un sector de la ciudad de San Sebastián. B) Nube de puntos clasificada en la que se pueden distinguir los puntos pertenecientes al suelo (coloreados en magenta) del resto de puntos de vegetación y edificaciones (coloreados en verde y en rojo, respectivamente).

Comparado con los métodos clásicos de adquisición manual, donde la interpretación y abstracción de

la topografía es realizada en el mismo momento en el que el operario toma el dato, los sistemas LiDAR realizan mediciones automáticas de la superficie sin ningún tipo de interpretación. Esto obliga a que sea necesaria una labor de post-procesado de los datos para extraer las mediciones que corresponden al suelo desnudo del conjunto de información registrada. Se trata de llevar a cabo un proceso de clasificación o “filtrado” que se basa en la utilización de algoritmos matemáticos o “filtros” (Kraus and Pfeifer 1998, Vosselman and Maas 2010, Polat and Uysal 2015) (ver Figura 6).

En el caso concreto de las aplicaciones forestales, el filtrado es un paso fundamental para obtener un MDE que permita calcular la altura real de los retornos láser sobre el terreno, de manera que a partir de ellos se pueda derivar una serie de variables explicativas íntimamente relacionadas con la estructura de la vegetación (Naesset 1997, Lloyd and Atkinson 2006, Nord-Larsen and Schumacher 2012, Maltamo *et al.* 2014).

De acuerdo con Silván-Cárdenas and Wang (2006), el valor de elevación de un retorno láser (Z) puede entenderse como una agregación de tres elementos: la elevación propia del terreno (t), la altura del objeto localizado sobre la superficie (f) y el error vertical (e), que puede deberse a varios factores (ecuación 2).

$$Z = t + f + e \quad (2)$$

Por lo tanto, el algoritmo de filtrado debe sopesar los componentes anteriores, no existiendo una solución única. Dependiendo del método de filtrado utilizado, el MDE resultante tendrá mayor o menor calidad (Sulaiman *et al.* 2010). De ahí que el desarrollo de métodos eficientes, efectivos y potentes para el filtrado del terreno sea actualmente un tema activo de investigación (Polat and Uysal 2015).

Aunque los algoritmos para extraer los puntos de suelo desnudo son cada vez más avanzados y automáticos (Axelsson 1999, Zhang *et al.* 2003, Silván-Cárdenas and Wang 2006, Mongus and Žalik 2012, Zhang and Lin 2013, Wang *et al.* 2015), siempre es necesario cierto grado de control manual a la hora de introducir los parámetros más adecuados (Baltsavias 1999a). La mayoría de estos algoritmos de filtrado funcionan bien en espacios relativamente llanos y con vegetación escasa o moderadamente densa. Sin embargo, el filtrado en áreas con terreno muy accidentado y/o discontinuo, pendientes pronunciadas y vegetación densa y adosada al suelo sigue siendo un desafío (Axelsson 1999, Vosselman 2000, Pingel *et al.* 2013). De acuerdo con Zhao *et al.* (2016), los algoritmos de filtrado existentes generalmente se testan en sitios específicos, por lo que su adaptabilidad a diferentes paisajes forestales es incierta y debe investigarse. Por lo tanto, es crucial desarrollar un algoritmo de filtrado que pueda hacer frente a diversos paisajes forestales, especialmente para regiones topográficas y ambientalmente complejas, como es el caso de la región Mediterránea.

La gran mayoría de los algoritmos de filtrado se basan en criterios geométricos y se pueden agrupar en las siguientes categorías según sus principios metodológicos (Sithole and Vosselman 2004, Meng *et al.* 2010): basados en superficies de interpolación (Kraus and Pfeifer 1998, Axelsson 1999, Evans and Hudak 2007, Kobler *et al.* 2007, Mongus and Žalik 2012, Hu *et al.* 2014), morfológicos (Kilian *et al.* 1996, Zhang *et al.* 2003, Chen *et al.* 2007, Li *et al.* 2013, Pingel *et al.* 2013) y los basados en la pendiente (Vosselman 2000, Roggero 2001, Sithole 2001).

Ante la diversidad de filtros implementados tanto en *software* de código abierto como propietario (Sulaiman *et al.* 2010), conviene conocer los fundamentos de cada uno de ellos para seleccionar aquel que sea más adecuado para el propósito del estudio (Meng *et al.* 2010). A nivel internacional destaca el estudio pionero realizado por Sithole and Vosselman (2004), patrocinado por la *International Society of Photogrammetry and Remote Sensing* (ISPRS), en el cual se testaron ocho algoritmos distintos para filtrar nubes de puntos con diferentes densidades y en paisajes tanto rurales como urbanos. El análisis

cualitativo y cuantitativo reveló que, mientras la mayor parte de los algoritmos funcionaban bien en áreas de baja complejidad (terreno suave, pequeños edificios, vegetación escasa) y con una alta proporción de puntos de suelo, el resultado fue significativamente peor en paisajes complejos (áreas urbanas, altas variaciones de elevación del terreno). En lo que respecta a la vegetación, todos los algoritmos funcionaron bien ante vegetación de porte alto. Tres de los ocho algoritmos analizados proporcionaron un filtrado eficaz de la vegetación ubicada en desniveles topográficos y de bajo porte. El estudio concluyó que los filtros de densificación progresiva de una superficie TIN (*Triangular Irregular Network*) y los basados en la interpolación de superficies, produjeron los mejores resultados (ver Capítulo 2 para más información sobre el funcionamiento de este tipo de filtros).

Tras el proceso de filtrado, para convertir los puntos de suelo en una superficie continua, hay que utilizar un método de interpolación apropiado (Zhang and Whitman 2005), que generalmente da lugar a dos estructuras de datos diferenciadas: TIN o ráster compuesto por una malla o matriz regular de píxeles o celdas. Este último es el más habitual por su aplicabilidad en diversos análisis y representa en 2,5 dimensiones la altura (Z) a través de la ecuación 3:

$$Z = F(x, y) \quad (3)$$

Donde el valor de (Z) en un área local determinada (píxel) depende de su correspondiente localización bidimensional (x, y) (Vosselman and Maas 2010).

Actualmente existen diversas funciones matemáticas para ajustar una superficie continua a partir de puntos de elevación muestreados sobre el terreno, tales como la distancia inversa ponderada (*IDW-Inverse Distance Weighting*), las superficies de tendencia polinómica local y global, el *kriging* o los algoritmos especializados como ANUDEM (Kienzle 2004, Hutchinson *et al.* 2011), cada uno de los cuales con sus propios parámetros, que el usuario debe definir para obtener un MDE lo más preciso posible.

El método de interpolación utilizado, junto con otros factores como la morfología del terreno, la precisión y la densidad de muestreo de la nube de puntos, y la resolución espacial de la matriz, influyen en la calidad de los MDE, tal y como evidencian números estudios (Caruso and Quarta 1998, Behan 2000, Gong *et al.* 2000, Thompson *et al.* 2001, Lloyd and Atkinson 2002, Kienzle 2004, Smith *et al.* 2004, Aguilar *et al.* 2005, Anderson *et al.* 2006, Polat *et al.* 2015). En consecuencia, las variables derivadas de los MDE y los análisis basados en estos modelos pueden verse afectados por un cierto grado de error (Kienzle 2004, Smith *et al.* 2004), por lo que interesa conocer la magnitud del error y sus características. La descripción de la calidad de un MDE puede llevarse a cabo de dos formas (Vosselman and Maas 2010):

- a) Según la calidad de los datos de entrada: precisión de las mediciones, densidad de puntos y procesamiento de la nube de puntos.
- b) Según la calidad del modelo *per se*, que puede dividirse en calidad interna y externa. La primera se basa en el análisis de la diferencia entre el dato de entrada y el estimado por el MDE a través de estadísticos como el error de la raíz cuadrada de la media (RMSE), el cual puede espacializarse para crear mapas que representen la precisión y así poder detectar las áreas más problemáticas. La segunda describe la calidad del MDE con respecto a unas mediciones externas de control, las cuales no habrán sido utilizadas en el proceso de generación del MDE y deberán tener una precisión superior a la de los datos de entrada para construir el MDE. Por ejemplo, la Sociedad Estadounidense de Fotogrametría y Teledetección (*American Society of Photogrammetry and Remote Sensing-ASPRS*) recomienda una exactitud vertical al menos tres veces superior a la del dato LiDAR. Es decir, la calidad externa considera tanto la de los datos de entrada, como la del proceso de modelado.

Estudios como el de Pfeifer *et al.* (2001) ya evidenciaron la estrecha relación entre el valor de RMSE obtenido y las características de la superficie, encontrando mayores errores en las zonas cubiertas por vegetación, frente a las zonas abiertas y desprovistas de ésta. La investigación desarrollada por Bater and Coops (2009), centrada en el análisis de siete métodos de interpolación (lineal, cuantiles, vecino natural, *spline* regularizado y de tensión, ANUDEM e IDW), consideró tres resoluciones espaciales del modelo (0,5, 1,0 y 1,5 m) y diferentes escenarios. Los resultados pusieron de manifiesto que la resolución espacial de 0,5 fue la más precisa. Los interpoladores vecino natural y lineal fueron los métodos más conservadores, con un rango de error más pequeño. Por el contrario, métodos como el IDW presentaron *outliers* que excedían ± 6 m, fundamentalmente debido a la complejidad en la parametrización del algoritmo. Los errores absolutos medios en zonas forestales de menor porte oscilaron entre 0,18 y 0,23 m, mientras que los errores medios absolutos en la categoría de bosque maduro fluctuaron entre 0,10-0,13 m.

Como describe Watt *et al.* (2013), la calidad general de los MDE derivados de datos LiDAR está influenciada por una serie de errores que pueden agruparse en cuatro categorías:

1. Errores causados por el propio sistema de medición (sistema láser y GPS).
2. Errores debidos a las características de los datos (retornos registrados, densidad de puntos, altura de vuelo, ángulo de escaneo). Conviene remarcar que, generalmente, los MDE para aplicaciones forestales a gran escala se obtienen a partir de vuelos LiDAR a 1-2 km sobre el suelo con diámetros de huella en torno a 0,5 m para garantizar la rentabilidad de la misión, dando como resultado bajas densidades de puntos (Maguya *et al.* 2013). Hyypä *et al.* (2005) señalan que el uso del último eco garantiza una mayor precisión del MDE, mientras que Su and Bork (2006) afirman que deberían utilizarse los puntos capturados con un ángulo igual o inferior a 15° del nadir, porque producen mejores resultados (ver Figura 7).
3. Errores creados durante el procesamiento de datos. De acuerdo con Guo *et al.* (2010), la disponibilidad de puntos de suelo desnudo para crear el MDE influye particularmente en aquellos considerados de alta resolución espacial (0,5 y 1 m). Según Behan (2000), la resolución espacial del MDE para cualquier método de interpolación debe estar lo más cerca posible, o ser ligeramente menor, al espaciamiento de puntos original. Las investigaciones coinciden en que las áreas de menor complejidad son las menos sensibles a la resolución espacial del modelo (Anderson *et al.* 2006). En última instancia, la elección del algoritmo óptimo de interpolación debe ser tomada por el usuario en función del tipo de área de estudio y de los datos de partida.
4. Errores debidos a las características de los objetos presentes en la escena. Se ha demostrado que la morfología del terreno, la estación del año, la densidad del arbolado y del sotobosque, y las especies forestales son factores importantes (Naesset 2002, Naesset and Okland 2002, Reutebuch *et al.* 2003). La influencia de la vegetación ha sido evaluada en diversos trabajos, como los de Reutebuch *et al.* (2003), Hopkinson *et al.* (2004), Su and Bork (2006), Bater and Coops (2009), Stereńczak and Kozak (2011) y Stereńczak *et al.* (2016). La estación del año repercute en la precisión del MDE, especialmente en bosques con un único estrato vegetal. Por ejemplo, la captura de datos LiDAR en la temporada de mayor vigor vegetal (presencia de hojas en las ramas de los árboles, crecimiento de los pastizales) puede incrementar el error en el MDE de 3 a 9 cm (Hyypä *et al.* 2005, Stereńczak and Kozak 2011). Las especies de árboles también juegan un papel importante, tal como señalaron Stereńczak and Kozak (2011) en su estudio. Los mayores errores los encontraron en zonas ocupadas por alerces o alisos frente a las cubiertas por pino o roble. En áreas abiertas con pastizales, los valores de error se correlacionan positivamente con la altura de la vegetación (Hopkinson *et al.* 2004). Los ambientes forestales particularmente complejos para la interpolación corresponden con áreas de vegetación densa,

tales como bosques jóvenes y matorrales, que pueden generar errores de varios metros de altura (Hopkinson *et al.* 2004, Bater and Coops 2009). Tales errores surgen porque difícilmente los pulsos láser pueden penetrar la densa estructura vegetal que oculta el suelo y, por consiguiente, la proporción de retornos pertenecientes al suelo en estas zonas es reducida (Watt *et al.* 2013). Las conclusiones sobre la influencia de la pendiente en la precisión del MDE son comunes en numerosos estudios (Aguilar *et al.* 2005, Hodgson *et al.* 2005, Hyyppä *et al.* 2005, Sterenczak *et al.* 2016), en los cuales, los valores de RMSE y sesgo se incrementan a medida que lo hace la pendiente topográfica.

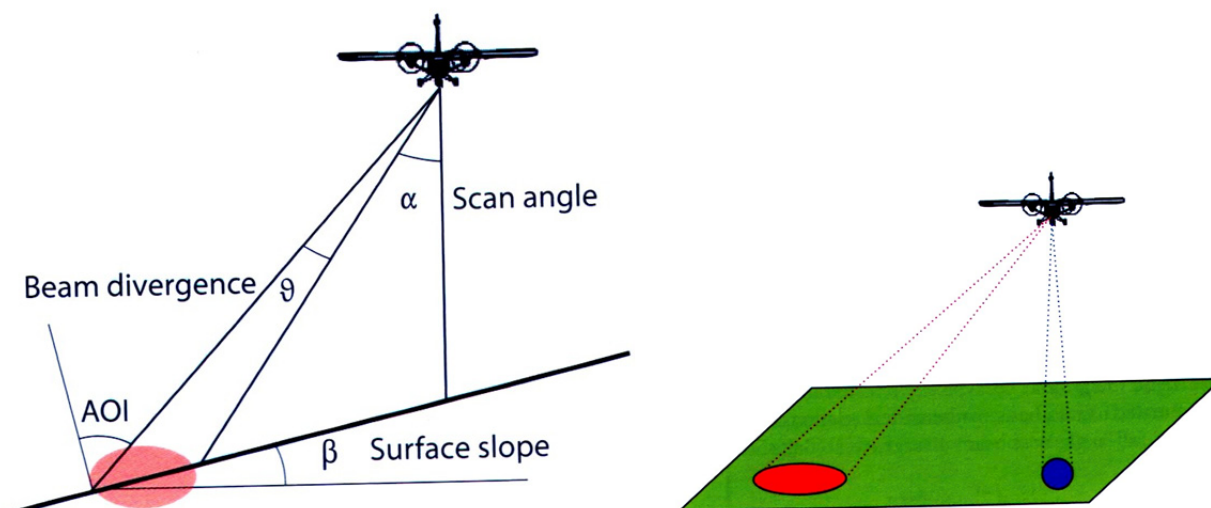


Figura 7. Ejemplo de distorsión de la huella láser sobre el terreno. Durante el proceso de exploración, el ángulo de incidencia (AOI) está continuamente cambiando debido a que también cambia el ángulo de escaneo, α , y la pendiente local del terreno, β . La combinación de estos dos elementos produce una elongación de la huella sobre el terreno, lo que puede dar lugar a un cierto error de medición de la distancia que separa el objeto del sensor. Si la huella se expande en exceso, la energía del pulso láser se reduce, y es posible que incluso no tenga suficiente potencia para regresar al sensor. Tomada de (Renslow 2013).

Parece evidente, por tanto, que los datos de elevación obtenidos con LiDAR son menos costosos de obtener y permiten la generación de modelos detallados de la superficie terrestre, lo que le ha valido a esta tecnología el reconocimiento. Al mismo tiempo, aparecen nuevas cuestiones relacionadas con el procesamiento de los datos y la calidad de los productos derivados.

Por otro lado, el LiDAR se ha utilizado como una técnica alternativa para derivar variables forestales de inventario y para analizar la estructura vertical y horizontal de la vegetación. La información espectral y textural que se puede obtener de las imágenes de teledetección multi e hiperspectrales (Liu *et al.* 2008, Gómez *et al.* 2012) ha sido explotada para estudiar, por ejemplo, la cobertura de arbolado (Céline *et al.* 2013), la densidad foliar (Glenn *et al.* 2008), la biomasa y la estructura del dosel (Wang *et al.* 2005, Castillo-Santiago *et al.* 2010, Pfeifer *et al.* 2012, Beguet *et al.* 2013). Sin embargo, todos estos trabajos presentan la misma perspectiva: la obtención de información estructural a nivel horizontal, dado que los sensores ópticos no registran directamente ningún tipo de información sobre la altura, ni sobre los estratos múltiples de vegetación. Las fotografías aéreas también se han utilizado para tratar de estimar parámetros estructurales relacionados con la altura del arbolado, el diámetro de copas, el cerramiento del dosel (vinculado

con la densidad de fustes y el volumen de la masa), etc., pero las imprecisiones son altas si se comparan con los sensores de teledetección activa (Campbell 2006), como el LiDAR o el SAR (*Synthetic Aperture Radar*). Estos últimos operan en el rango de las microondas y ofrecen observaciones mundiales con alta resolución espacial y temporal. Además, las longitudes de onda más largas permiten observaciones de la superficie terrestre independientemente de la cobertura de nubes y de la iluminación. Más concretamente, los sistemas Interferométricos SAR (InSAR) facilitan la adquisición de información tridimensional de la superficie y constituyen una base apropiada para la investigación de la estructura forestal en un contexto global. Dependiendo de la longitud de onda y, por tanto, de la capacidad de penetración de la señal, estos sistemas son sensibles a la distribución de las hojas, ramas y troncos, y proporcionan información sobre los estratos verticales (Varghese *et al.* 2016). Comparados con el LiDAR, los datos InSAR son ventajosos por cuanto están disponibles a escala mundial y su adquisición es más rentable; sin embargo, pocos estudios los utilizan debido a su complejidad frente a los primeros (Abdullahi *et al.* 2017).

Actualmente existe una gran cantidad de literatura científica sobre las posibilidades de aplicación del LiDAR aeroportado de pulsos discretos en el sector forestal, en particular sobre la estimación de biomasa y de otras variables estructurales propias de los inventarios forestales (Wulder *et al.* 2008, Maltamo *et al.* 2014), que son de interés para la gestión sostenible de los bosques (Guerra-Hernández *et al.* 2016). Dependiendo de la densidad de puntos disponibles, del objetivo y de la precisión que se quiera conseguir (Yu *et al.*, 2010), pueden llevarse a cabo dos aproximaciones diferentes para derivar información forestal: el enfoque basado en áreas (ABA-*area-based approach*) y el enfoque basado en árbol individual (ITB-*individual tree-based approach*) (Reutebuch *et al.* 2005, Yu *et al.* 2011, Latifi *et al.* 2015).

En el enfoque ABA, las características de la masa forestal de un área determinada son estimadas estableciendo relaciones empíricas (modelos) entre los datos LiDAR y las variables medidas sobre el terreno (Reutebuch *et al.* 2005). Este enfoque fue ideado originalmente por Naesset (1997) y también se conoce como el procedimiento de inventario de masas en dos fases o inventario forestal de doble muestreo (Naesset 2002). En la primera fase se obtienen las relaciones empíricas entre los retornos láser (variables descriptivas de la distribución en alturas de los puntos) y los datos medidos sobre el terreno (por ejemplo, altura, área basimétrica, volumen, biomasa aérea, etc.) a nivel de parcelas de inventario, generalmente circulares. En la segunda fase, estas relaciones se utilizan para ajustar modelos predictivos que puedan ser extrapolados a toda el área de estudio (Rosette *et al.* 2012). Los resultados de las investigaciones llevadas a cabo siguiendo este enfoque han demostrado la utilidad para estimar variables tales como la densidad de árboles (Naesset 2002, Lindberg and Hollaus 2012), la altura media (Naesset 2002, Coops *et al.* 2007, González-Ferreiro *et al.* 2012, Smreček and Danihelová 2013), el diámetro medio (Naesset 2002), el área basal, el volumen (Naesset 1997, Latifi *et al.* 2010, Estornell *et al.* 2012, Alberti *et al.* 2013), la biomasa aérea (Naesset and Gobakken 2008, Ene *et al.* 2012, Kankare *et al.* 2013) y las reservas de carbono (Patenaude *et al.* 2004, García *et al.* 2010, Stephens *et al.* 2012). Además, estos métodos ABA se han adaptado para caracterizar combustibles forestales (Riaño *et al.* 2003) y derivar el índice de área foliar (Riaño *et al.* 2004b, Morsdorf *et al.* 2006, Richardson *et al.* 2009). La ventaja de los métodos ABA radica en el hecho de que son aplicables cuando las densidades de puntos LiDAR son bajas. Por el contrario, requieren de más mediciones en campo (Yu *et al.* 2010) y los modelos derivados son aplicables a nivel local, es decir, son específicos para ciertas áreas, tipos de masas forestales y características del vuelo (altura de la aeronave, densidad de puntos, ángulo de escaneo, patrón de escaneo, etc.) (Rosette *et al.* 2012).

El enfoque ITB se basa en la delineación de la copa (diámetro y superficie), la medida de la altura y la detección de los ápices de cada árbol para estimar, sobre esa base, otras variables como el

diámetro a la altura del pecho, el área basimétrica, el volumen, la biomasa, el carbono, etc. Al igual que el enfoque ABA, el ITB también requiere de un conjunto de mediciones directas sobre el terreno en parcelas de muestreo dentro de la zona de estudio. Sin embargo, este enfoque funciona relativamente peor en masas complejas donde los algoritmos de segmentación de copas producen frecuentemente errores de comisión y omisión (Guerra-Hernández *et al.* 2016). Además, es necesario contar con nubes de una densidad de, al menos, 4-5 puntos/m² (Andersen *et al.* 2006). Por tanto, la aproximación ABA es una alternativa menos costosa y comúnmente utilizada para generar mapas de atributos forestales en una amplia gama de tipos de bosque: templado (Hall *et al.* 2005), boreal (Næsset and Gobakken 2008), atlántico (Gonçalves-Seco *et al.* 2011) y mediterráneo (González-Olabarria *et al.* 2012). Además, los resultados demuestran las altas correlaciones que existen entre las variables derivadas de los retornos láser y las variables de inventario forestal (Maltamo *et al.* 2014).

La estructura de la vegetación describe el tamaño, la forma y la distribución de los componentes del dosel, como son los fustes, las ramas y las hojas. Estas características no son estáticas; cambian según la fenología de la especie, el estrés ambiental, la dinámica del ecosistema y las perturbaciones naturales o antrópicas. En este sentido, los incendios forestales son considerados uno de los factores de perturbación más importantes en los ecosistemas naturales (Lasaponara *et al.* 2006). Particularmente, en la Cuenca Mediterránea, el fuego ha sido históricamente un elemento natural modelador del paisaje, con el cual las sociedades han aprendido a coexistir hasta convertirlo en herramienta de gestión tradicional. Sin embargo, en las últimas décadas, la recurrencia, magnitud y severidad de los incendios se ha incrementado, siendo una de las principales causas el aumento de la combustibilidad (cantidad, superficie, volumen, continuidad espacial y materia muerta) de las masas vegetales, como consecuencia de los cambios en la estructura agraria y en los usos de suelo. A ello hay que sumar el gran desafío ambiental y socioeconómico que implica el cambio climático, con condiciones meteorológicas extremas y periodos prolongados de sequía (Solomon *et al.* 2007, González-De Vega *et al.* 2016). Por otro lado, también influye el tipo de especie forestal presente en el ecosistema. Los bosques de *Pinus halepensis* (en adelante *P. halepensis*) son las formaciones de pinos más abundantes de la Cuenca Mediterránea. Están muy bien adaptados a las condiciones de sequía y a unos suelos pobres; sin embargo, por su naturaleza (alta densidad, presencia de ramas a lo largo del fuste y pocos tratamientos silvícolas) son muy inflamables (Pausas *et al.* 2008). Por todo ello, el régimen de incendios de estos ecosistemas se caracteriza por la alta intensidad, con un intervalo de ocurrencia promedio de aproximadamente 30 a 50 años. A pesar de la alta resiliencia post-fuego, la regeneración de estos bosques de pinos puede fallar cuando los intervalos de tiempo entre los incendios son más cortos que el tiempo requerido para acumular un banco de semillas suficiente (Tanase *et al.* 2015, González-De Vega *et al.* 2016).

Desde el punto de vista de la gestión forestal, la evaluación y cartografía de la severidad post-incendio proporciona información fundamental para priorizar la aplicación de tratamientos y medidas que minimicen los efectos del fuego y reduzcan el tiempo de recuperación del ecosistema afectado (Montorio *et al.* 2014). Además, la severidad del fuego puede usarse para predecir la respuesta del ecosistema en términos de regeneración vegetal y de dinámica hidro-geomorfológica (Tanase *et al.* 2015).

De forma general, el concepto de severidad hace referencia a la magnitud del cambio ecológico causado por el fuego (Key and Benson 2006, Lentile *et al.* 2006). No obstante, en la bibliografía consultada existen ciertos matices respecto de este término que conviene aclarar. Por un lado, la evaluación de los efectos inmediatos del fuego se conoce como severidad del fuego (*fire severity* en inglés), mientras que la evaluación de la respuesta ecológica o ambiental se denomina severidad del incendio (*burn severity*), la cual es función de la primera (Chuvieco, 2009, Montorio *et al.* 2014, Tanase *et al.* 2015).

La severidad del fuego puede ser evaluada a partir de la observación directa en el campo para reconocer los niveles del daño causado sobre el medio. Para ello se utilizan un conjunto limitado de indicadores (Kasischke *et al.* 2008) que pueden estar orientados a la evaluación del impacto en la vegetación (por ejemplo, la mortalidad de los árboles, los daños en la copa y el fuste, la altura de la quemadura) y/o al análisis del suelo quemado (profundidad de la capa de materia orgánica, cubierta de hojarasca, repelencia al agua, temperatura del suelo y humedad, etc.), sobre todo para conocer los procesos posteriores al fuego, como la germinación y supervivencia de las plántulas (Tanase *et al.* 2015). Más frecuentemente suelen utilizarse índices sintéticos que combinan diferentes métricas, como el *Composite Burn Index* (CBI) (Key and Benson 2003) u otras variantes de éste como el GeoCBI (De Santis and Chuvieco 2009). Mediante el CBI se examinan visualmente las condiciones de la vegetación y se registra el grado de cambio con respecto al supuesto estado previo al fuego, a nivel de sotobosque, dosel arbóreo y a nivel global de la parcela de campo evaluada. Las técnicas de teledetección basadas en los cambios en la reflectividad de las áreas quemadas –debidos al proceso de combustión que implica la eliminación de la vegetación, la exposición del suelo, los cambios en el contenido de humedad y la aparición de nuevas cubiertas– proporcionan información que puede relacionarse a través de índices espectrales por combinación de bandas reflectivas con el dato de campo, para estimar y cuantificar de forma más objetiva la severidad (Chuvieco 2009). Entre los más conocidos destacan el *Normalized Difference Vegetation Index* (Rouse *et al.* 1974), el *Normalized Burn Ratio* (García and Caselles 1991), el *delta Normalized Burn Ratio* (Key and Benson 2006), el *Relative delta Normalized Burn Ratio* (Miller and Thode 2007), el *SWIR-MIR Index* (Veraverbeke *et al.* 2012) y el *Relativized Burn Ratio* (Parks *et al.* 2014).

Habitualmente, las estimaciones de campo y los datos de teledetección se han relacionado empleando modelos de regresión lineal, cuadrática y exponencial, aunque han cobrado importancia en estos últimos años los modelos no paramétricos, como alternativa al incumplimiento del supuesto de normalidad en la distribución de los datos de la muestra. Por ejemplo, la relación entre el dNBR y el CBI ha proporcionado una detección precisa de la severidad en la mayoría de los ecosistemas, siendo generalmente altos los coeficientes de determinación obtenidos ($R^2 > 0,75$) (Tanase *et al.* 2011). No obstante, los índices espectrales suelen producir peores resultados (Tanase *et al.*, 2011) a la hora de identificar niveles de severidad moderados y bajos, ya que se utiliza la respuesta espectral del menor contenido de humedad de la vegetación y el suelo, el aumento de la exposición del sustrato y la presencia de cenizas, que son relativamente estables para niveles altos de severidad (Chuvieco 2009). Además, los índices basados en información óptica son también sensibles a la fenología de las plantas, a la elevación solar (Verbyla *et al.* 2008) y a la cobertura de nubes, lo que disminuye su utilidad en determinados ambientes (Tanase *et al.* 2015).

Estudios recientes han demostrado que la severidad del fuego, estimada a través del índice CBI, puede relacionarse empíricamente con datos procedentes de sensores activos como el SAR o el LiDAR (Tanase *et al.* 2010, Tanase *et al.* 2014). La ventaja de éstos frente a los ópticos es que proporcionan una medida directa de la estructura de la vegetación. En el caso del LiDAR, su uso para estimar la severidad es todavía incipiente. Sobre todo se ha empleado en combinación con series temporales de imágenes Landsat, que conjuntamente permiten detectar las alteraciones y cuantificar el impacto del fuego sobre la estructura de la vegetación (Wulder *et al.* 2009, Kane *et al.* 2013, Kane *et al.* 2014, Bolton *et al.* 2015).

Otro aspecto a tener en cuenta a la hora de valorar la amenaza del fuego en un territorio es la actividad que en él se desarrolla. La evaluación del riesgo de incendio, que es clave para la prevención y la planificación previa a la extinción, integra la estimación de la ignición, las condiciones de propagación y la valoración de la vulnerabilidad. En este sentido, la cartografía de tipos de combustible se revela muy importante para la estimación de la ignición y de las condiciones de propagación y, por ende,

para la valoración del peligro de incendio forestal.

Los mapas de modelos de combustible que tradicionalmente se han elaborado en España se han basado en el Mapa Forestal de España escala 1:50.000, que tiene una periodicidad de renovación excesivamente dilatada para la dinámica de los ecosistemas mediterráneos, dando lugar a resultados groseros e imprecisos para la gestión eficaz de los incendios a escala local. Como cualquier simplificación de la realidad, los sistemas de clasificación de tipos de combustibles presentan ciertas limitaciones, ya que son únicamente aplicables en localizaciones geográficas con características similares a aquellas en las que fueron desarrolladas (Arroyo *et al.* 2008). Además, la cartografía de tipos de combustible suele requerir de trabajo de campo ante la inexistencia de un método automático y fiable para la identificación de los combustibles desde gabinete.

Los primeros trabajos para cartografiar los tipos de combustible se iniciaron a mediados de la década de 1930 en EE.UU., cuando todavía no se hablaba de “tipos de combustible”, sino de tipos de vegetación o cobertura y únicamente se contaba con el trabajo de campo a pie. Con la llegada de la fotografía aérea a comienzos de 1940 se dio un salto cualitativo, al poder llevar a cabo técnicas de fotointerpretación para identificar el combustible y así reducir el intenso trabajo de campo (Chuvieco and Martín 2004). La aparición de nuevos sensores remotos en la década de los setenta del siglo XX fue lo que verdaderamente revolucionó la cartografía de modelos de combustibles, coincidiendo con la publicación de Albin (1976) y del primer trabajo de Rothermel (1972) sobre el comportamiento del fuego. La disponibilidad de imágenes en cortos intervalos de tiempo, la reducción de costes frente a la fotografía aérea, la interrelación entre el dato observado en el campo y la información de la imagen, la toma de datos homogénea para grandes áreas, la captura de información no visible del espectro electromagnético (infrarrojo cercano, infrarrojo medio, infrarrojo térmico y micro-ondas), la naturaleza digital de los datos de teledetección y la posibilidad de incluir otras variables espaciales en el análisis, son algunas de las ventajas frente a los métodos tradicionales para la cartografía de modelos de combustibles (Chuvieco and Martín 2004). Por todos estos motivos, las imágenes ópticas multiespectrales captadas con sensores de tipo pasivo han sido habitualmente utilizadas en la cartografía de tipos de combustible a escala local y regional, como un procedimiento alternativo frente a los costosos métodos de identificación a partir de fotografías aéreas o directamente en campo (Burgan *et al.* 1998, Tian *et al.* 2005, Lasaponara and Lanorte 2007). Con respecto a los sensores más empleados para este propósito, destacan en primer lugar los embarcados en los satélites Landsat, debido a su media-alta resolución espacial (30 m), a su resolución espectral (cuenta con una banda en el infrarrojo próximo y dos en el infrarrojo medio de onda corta, muy útiles para diferenciar la vegetación) y a una resolución temporal que supera los requerimientos de este tipo de cartografía. En menor medida se ha utilizado el sensor NOAA-AVHRR (1 km de resolución espacial), los datos obtenidos con SPOT (20-10 m) y los sensores hiperespectrales aeroportados, como AVIRIS, que proporcionan una altísima resolución espectral (210 bandas entre 0,4 y 2,4 μm y 20 m de resolución espacial) muy útil para discriminar las cubiertas (Chuvieco and Martín 2004).

Sin embargo, estos sistemas de teledetección pasiva presentan ciertas limitaciones, dada su incapacidad para detectar el combustible localizado bajo el dosel arbóreo (sotobosque y estratos inferiores (Vosselman and Maas 2010)), que es de crucial importancia para determinar el comportamiento del fuego. Además, la reflectividad que es captada por estos sensores –que debe someterse a correcciones radiométricas del efecto topográfico y atmosférico– no puede relacionarse con variables estructurales como la altura de la vegetación, que es un parámetro imprescindible en la clasificación de tipos de combustible (Riaño *et al.* 2002). A este respecto, los sensores de teledetección activos sí que proporcionan información sobre la estructura vertical de la vegetación (Chuvieco 2010). En el caso del SAR, existen relaciones estadísticas significativas entre el tamaño de los componentes de la cubierta vegetal y el coeficiente de retrodispersión obtenido, siendo también posible el uso de

técnicas polarimétricas e interferométricas para modelizar esta cuestión (Chuvieco and Martín 2004, Chuvieco 2010). No obstante, las reflexiones en la longitud de onda de las microondas se ven afectadas por el efecto topográfico, que es especialmente complejo de corregir en zonas de orografía montañosa, lo que implica una reducción en la precisión frente al LiDAR. En este caso, los escáner láser se han utilizado con éxito para la obtención de cartografía de modelos de combustible (Riaño *et al.* 2007, Mutlu *et al.* 2008) y el modelado de otras variables necesarias para la predicción del comportamiento del fuego: la densidad aparente del dosel de copas (Riaño *et al.* 2004a, Andersen *et al.* 2005, Erdody and Moskal 2010), la altura de la base del dosel (Riaño *et al.* 2003, Popescu and Zhao 2008, Maguya *et al.* 2015) y la cubierta de copas (Hall *et al.* 2005). Otros estudios se han centrado en elementos específicos de los combustibles, tales como la densidad de los combustibles de escalera que conectan el suelo con la copa de los árboles (Kramer *et al.* 2014, Platt 2014).

La clasificación digital de bandas de información ha sido la técnica más utilizada frente a la interpretación visual de las imágenes, sobre todo la clasificación supervisada con el algoritmo de máxima probabilidad, la clasificación no supervisada, los clústeres o conglomerados y la clasificación mixta, aunque en estos últimos años han cobrado importancia otros métodos basados en redes neuronales, árboles de decisión, etc. Particularmente se observa una tendencia orientada a la combinación de datos LiDAR e imágenes ópticas y SAR, para aprovechar las ventajas propias de cada fuente de información, evidenciándose una mejora en la precisión de la cartografía de combustibles (Mutlu *et al.* 2008, Koetz *et al.* 2008, García *et al.* 2011, Jakubowski *et al.* 2013).

1.2 Justificación

La tecnología LiDAR aeroportada de pulsos discretos se ha consolidado como un medio para adquirir información topográfica georreferenciada y precisa. Más allá de sus aplicaciones primarias, como la generación de MDE, ha demostrado ser útil y adecuada para otros análisis que tienen que ver con los espacios forestales, reconociéndose su valor frente al uso exclusivo de imágenes ópticas, carentes de información tridimensional sobre los elementos de la escena. Puede afirmarse que ha habido un cambio progresivo en el enfoque de las investigaciones: desde el uso del LiDAR como un simple método de captura de datos, al desarrollo de protocolos y metodologías analíticas eficaces.

Los estudios para determinar cuáles son los mejores procedimientos para el tratamiento de los datos LiDAR pueden ayudar al usuario a adoptar el más adecuado. Las nubes de puntos deben ser clasificadas utilizando algún algoritmo específico que hay que parametrizar y que depende de las características intrínsecas de los datos LiDAR y de la escena. Son pocos los estudios que comparan diferentes algoritmos de filtrado, y mucho menos utilizando los nuevos datos del PNOA, a lo que hay que sumar una amplia variedad de métodos implementados en *software* libre y propietario. Además, hay que considerar que el error en el filtrado está condicionado, no solo por la idoneidad del algoritmo elegido, sino también por las características asociadas a los datos LiDAR adquiridos (densidad y espaciado de los puntos fundamentalmente) y por la complejidad de la superficie (vegetación en pendientes, discontinuidades topográficas, etc.).

Continuando con esta perspectiva que incide en la importancia de tratar exhaustivamente los datos LiDAR, hay que considerar la elección del método de interpolación para generar MDE en espacios forestales. Tal como se ha puesto de manifiesto en el apartado anterior, estudios previos han demostrado que la precisión de los MDE no solo está influenciada por los procesos que implican la manipulación de los puntos (método de filtrado y de interpolación elegidos), sino también por una serie de factores externos, tales como la variabilidad topográfica, la densidad de retornos de suelo, la complejidad estructural del dosel vegetal y la resolución espacial, que conviene analizar. Por estas razones, conocer la calidad de un MDE generado con datos LiDAR es esencial, ya que podrá condicionar los resultados de las técnicas de modelado espacial aplicadas.

Por otro lado, la gestión y ordenación de montes implica, en muchas ocasiones, la estimación de variables dasométricas que tradicionalmente se han obtenido mediante un muestreo sistemático por parcelas, requiriendo una gran inversión económica e intenso trabajo de campo. En este contexto, es necesario evaluar la idoneidad de los datos públicos LiDAR-PNOA para estimar este tipo de variables mediante modelos de regresión que relacionen adecuadamente la información aportada por el LiDAR con la medida directamente en el campo. Más concretamente, la obtención de cartografía de biomasa forestal ha adquirido un interés creciente en los últimos años debido al papel que juegan los bosques en el ciclo del carbono y su importancia como recurso disponible para la producción de bioenergía. Numerosos estudios han correlacionado la biomasa y la respuesta espectral de la vegetación utilizando sensores ópticos pasivos; sin embargo, son susceptibles de sufrir problemas de saturación cuando la biomasa aérea es elevada, lo cual no afecta a los modelos de estimación realizados con la información estructural de las nubes de puntos LiDAR. Dichos modelos son necesarios para poder cuantificar el carbono retenido por la vegetación y el CO₂ emitido a la atmósfera.

En el contexto de los incendios forestales, el carácter tridimensional de los datos LiDAR posibilita estudiar los cambios estructurales ocurridos en la vegetación tras el fuego, por lo que se pueden realizar nuevas aportaciones en este ámbito que sirvan para apoyar las tareas de restauración del área quemada y proporcionen información sobre la respuesta que presumiblemente tendrá el ecosistema a los efectos del fuego. En este sentido, el trabajo de campo es clave para relacionar la severidad del fuego estimada mediante un índice visual y las variables derivadas de los retornos LiDAR. En esta misma línea, la cartografía de los tipos de combustible forestal es esencial para la estimación tanto de la ignición como de las condiciones de propagación del fuego en caso de incendio, además de ser muy útil para la gestión forestal. Los estudios basados en técnicas de clasificación digital de imágenes de teledetección han demostrado cómo la combinación de información proveniente de sensores pasivos y activos mejora los resultados que se obtienen con cada uno de estos tipos de datos por separado. Es por ello que la investigación debe caminar hacia el aprovechamiento de las bondades que presentan los diferentes tipos de sensores (pasivos y activos), con el fin de plantear metodologías integradoras.

Uno de los aspectos que se considera de mayor interés en esta investigación es el área de estudio, caracterizada por un ecosistema mediterráneo continentalizado, que es sinónimo de diversidad desde el punto de vista geográfico. Posiblemente sea ésta la razón por la que su estudio es complejo y los resultados que se obtienen implican una menor capacidad de generalización.

En definitiva, la presente tesis doctoral trata de enriquecer el acervo científico sobre el tratamiento de los datos LiDAR de baja densidad y arroja luz sobre la aplicación de técnicas para obtener cartografía de variables continuas (altura del dosel, biomasa, etc.) y discretas (por ejemplo, tipos de combustibles forestales), como instrumento para la gestión sostenible del territorio. Esto es especialmente relevante cuando se trata de vegetación mediterránea, la cual ha sufrido profundos cambios durante los últimos decenios, relacionados con alteraciones en el régimen de perturbaciones (incendios) y en el nivel de explotación (madera, leña, pastoreo, carboneo...).

1.3 Hipótesis y objetivos

La estimación de variables forestales continuas, como la altura media, el área basimétrica, el volumen, la densidad, la biomasa, etc., y discretas, como el tipo de combustible forestal, constituyen unas de las líneas de trabajo que más importancia han cobrado en estos últimos años con sensores activos de teledetección por las ventajas expuestas en el capítulo anterior. Además, la posibilidad de derivar variables o métricas de los retornos láser que impactan sobre la vegetación, convierte a esta tecnología en una herramienta verdaderamente interesante para relacionar esta información con observaciones realizadas directamente en el campo.

Sin embargo, existe una necesidad, desde el punto de vista metodológico (herramientas y técnicas), de estudiar los espacios forestales mediterráneos, caracterizados por ser pluriespecíficos, de escasa densidad forestal y alta heterogeneidad espacial, y por estar localizados en zonas de pendientes acusadas, lo cual puede dificultar la obtención de modelos robustos para estimar parámetros forestales diversos. Dicho interés responde a que en la literatura científica referente a la estimación de parámetros forestales continuos, abundan las experiencias en bosques con alta densidad y homogeneidad estructural y orográfica, donde el uso del LiDAR ha proporcionado resultados óptimos. Sin embargo, un menor número de investigaciones se ha centrado en medios más complejos, como los mediterráneos, y mucho menos utilizando datos del vuelo PNOA, que, con cubrimiento sistemático para todo el territorio nacional y financiado con fondos públicos, demandan su puesta en valor. Por todo ello, la hipótesis planteada supone la posibilidad de estimar de forma precisa parámetros estructurales en espacios forestales de formaciones vegetales heterogéneas y orografía compleja, empleando datos LiDAR de baja densidad de puntos con apoyo en trabajo de campo.

El objetivo general de esta tesis doctoral es el desarrollo de metodologías operativas para el procesamiento de los datos LiDAR-PNOA que, en último término, permitan el inventario de un conjunto de variables dasométricas de interés para los gestores del bosque, la valoración de los niveles de severidad post-fuego y la identificación de los combustibles forestales, todo ello en un ecosistema mediterráneo con rasgos semiáridos.

Para lograr este objetivo principal ha sido necesaria la consecución de una serie de objetivos específicos, los cuales se han abordado en cada una de las publicaciones (artículos científicos y comunicaciones en congresos) que componen el cuerpo de la tesis (ver Tabla 2):

1. Identificar el filtro o método de clasificación que mejor se adapta a las características de la nube de puntos LiDAR-PNOA en un ambiente forestal heterogéneo a través de una metodología semi-empírica, mediante la comparación de los implementados en *software* libre.
2. Analizar la idoneidad de diversos métodos de interpolación para crear MDE de alta resolución con los retornos de suelo desnudo, dado que son necesarios para la normalización de las alturas de la nube de puntos y, por ende, de la vegetación en el contexto de las aplicaciones forestales del LiDAR.
3. Estimar variables dasométricas en masas de *P. halepensis* Mill., así como el contenido de carbono de estas masas, empleando modelos de regresión lineal múltiple que relacionan las variables derivadas de los retornos láser y los datos de campo.
4. Analizar las relaciones que existen entre los indicadores de la severidad del fuego estimados en campo y las variables derivadas de las nubes de puntos mediante un ajuste de regresión logística para obtener una cartografía de cuatro áreas quemadas.
5. Desarrollar una metodología para obtener cartografía de tipos de combustible forestal mediante combinación de información LiDAR-PNOA e imágenes multiespectrales de alta resolución.

1.4 Estructura

El contenido de la tesis doctoral se ha estructurado en cuatro capítulos, más un apéndice, de acuerdo con los requisitos de la modalidad de tesis como compendio de publicaciones.

En este primer capítulo se presentan el estado de la cuestión, la justificación de la investigación y los objetivos. El segundo capítulo describe el área de estudio y los materiales utilizados, así como el conjunto de técnicas y métodos estadísticos que se incluyen en las publicaciones con el fin de mejorar la comprensión de los mismos. El tercer capítulo incluye la versión original de las publicaciones aceptadas y publicadas. Por último, el capítulo cuarto sintetiza las principales conclusiones que se derivan de los estudios realizados, que dan pie a líneas futuras de trabajo. En el apéndice se detallan el factor de

impacto de las revistas, las áreas temáticas correspondientes a las publicaciones que se recogen en la tesis y la justificación de la contribución del doctorando, al tratarse de trabajos publicados en coautoría.

En la Tabla 2 se presenta la secuencia de objetivos, las técnicas y métodos más importantes y las herramientas utilizadas, haciendo referencia a la publicación en la cual se recogen los principales resultados y conclusiones. Complementariamente, en la Figura 8 se muestra un esquema de la estructura general de la tesis, donde se puede apreciar su cohesión a través de la interconexión entre los estudios realizados (Capítulo 3), así como la relación entre las zonas de estudio (A, B y C) seleccionadas. En este sentido, conviene aclarar que, si bien el área de estudio se circunscribe al territorio de Aragón, dada la complejidad técnica de los tratamientos y el carácter experimental de las metodologías aplicadas, se han seleccionado tres zonas de particular interés en atención a los objetivos específicos de la investigación, las cuales se describirán en detalle en el Capítulo 2:

- Zona A: dos zonas test situadas en el valle medio del Ebro, en el entorno de los Montes de Zuera (Zaragoza).
- Zona B: el Centro de Adiestramiento (CENAD) “San Gregorio” y su entorno, ubicado al norte de la ciudad de Zaragoza.
- Zona C: cuatro incendios forestales localizados en el sector central de la Depresión del Ebro (Zuera y Jaulín) y en las estribaciones del Sistema Ibérico turolense (Aliaga y Los Olmos).

Tabla 2. Resumen de los objetivos específicos, técnicas y métodos, herramientas y publicaciones.

Objetivo específico 1	Técnicas y métodos	Software utilizado
Identificar el filtro que mejor se adapta a las características de la nube de puntos LiDAR-PNOA en un ambiente forestal heterogéneo.	<ul style="list-style-type: none"> - Clasificación de la nube de puntos con distintos filtros basados en superficies. - Validación cualitativa del filtrado. - Validación cuantitativa del filtrado (estadísticos de error de omisión y comisión, error total e índice Kappa) con apoyo en trabajo de campo. - Análisis del error considerando la influencia de la pendiente topográfica, el tipo de cubierta del suelo, la densidad de puntos y el ángulo de escaneo. 	<ul style="list-style-type: none"> - LAStools - FUSION v.3.30 - MCC-LiDAR v.2.1 - BCAL LiDAR Tools v.1.5.1 - ALDPAT v.1.0 - SPSS Statistics 20 - ArcGIS 10.1

Publicación

Montealegre, A.L., Lamelas, M.T., de la Riva, J. 2015. Comparison of open source LiDAR filtering algorithms in a Mediterranean forest environment. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 8: 4072-4085. doi:10.1109/JSTARS.2015.2436974.

Objetivo específico 2	Técnicas y métodos	Software utilizado
Analizar la idoneidad de diversos métodos de interpolación para crear MDE de alta resolución para la normalización de las alturas de la vegetación.	<ul style="list-style-type: none"> - Interpolación de la nube de puntos filtrada con distintos algoritmos. - Análisis de la calidad interna del MDE (estadísticos de error). - Análisis de la calidad externa del MDE con apoyo en trabajo de campo (estadísticos de error). - Análisis del error considerando la influencia de la pendiente topográfica, el tipo de cubierta del suelo, la densidad de puntos y el ratio de penetración de pulsos láser a través de la vegetación. - Análisis CART. 	<ul style="list-style-type: none"> - ArcGIS 10.1 - FUSION v.3.30 - SPSS Statistics 20

Publicación

Montealegre, A.L., Lamelas, M.T., de la Riva, J. 2015. Interpolation routines assessment in ALS-derived Digital Elevation Models for forestry applications. *Remote Sensing*, 7: 8631-8654. doi:10.3390/rs70708631.

Tabla 2. Continuación.

Objetivo específico 3	Técnicas y métodos	Software utilizado
Estimar variables dasométricas y del contenido de carbono en masas de <i>P. halepensis</i> Mill.	<ul style="list-style-type: none"> - Obtención de las variables dasométricas dependientes mediante trabajo de campo. - Obtención de las variables independientes derivadas de los retornos láser. - Análisis de correlación y de regresión lineal múltiple. - Validación cruzada dejando un dato fuera. - Generación de cartografía. 	<ul style="list-style-type: none"> - FUSION v.3.30 - MCC-LiDAR v.2.1 - Lenguaje y entorno de programación R - ArcGIS 10.1

Publicación

Montealegre, A.L., Lamelas, M.T., de la Riva, J., García-Martín, A., Escribano, F. 2016. Use of low point density ALS data to estimate stand-level structural variables in Mediterranean Aleppo pine forest. *Forestry*, 89 (4): 373-382. doi:10.1093/forestry/cpw008.

Montealegre, A.L., Lamelas, M.T., de la Riva, J., García-Martín, A., Escribano, F. 2015. Assessment of biomass and carbon content in a Mediterranean Aleppo pine forest using ALS data. In *Proceedings of the 1st Int. Electron. Conf. Remote Sens.*, 22 June–5 July 2015; Sciforum Electronic Conference Series, Vol. 1, d004. doi:10.3390/ecrs-1-d004.

Objetivo específico 4	Técnicas y métodos	Software utilizado
Analizar la relación entre el índice CBI y las estadísticas derivadas de las nubes de puntos para cartografiar la severidad.	<ul style="list-style-type: none"> - Estimación de la severidad post-incendio en campo (variable dependiente) mediante el protocolo CBI. - Obtención de las variables independientes derivadas de los retornos láser. - Análisis de correlación, de contraste entre muestras y de regresión logística. - Validación cruzada (matriz de confusión, índice Kappa y curvas de ROC). - Generación de cartografía. 	<ul style="list-style-type: none"> - FUSION v.3.30 - MCC-LiDAR v.2.1 - SPSS Statistics 20 - ArcGIS 10.1

Publicación

Montealegre, A.L., Lamelas, M.T., Tanase, M., de la Riva, J. 2014. Forest fire severity assessment using ALS data in a Mediterranean environment. *Remote Sensing*, 6: 4240-4265. doi:10.3390/rs6054240.

Objetivo específico 5	Técnicas y métodos	Software utilizado
Desarrollar una metodología para obtener cartografía de tipos de combustible forestal mediante combinación de la información LiDAR-PNOA e imágenes multiespectrales de alta resolución.	<ul style="list-style-type: none"> - Estimación de tipos de combustible en campo según el modelo Prometheus. - Procesamiento de los datos LiDAR y de las imágenes ópticas. - Clasificación digital (método paramétrico supervisado de máxima probabilidad). - Validación cruzada (matriz de confusión, índice Kappa). - Generación de cartografía. 	<ul style="list-style-type: none"> - FUSION v.3.30 - MCC-LiDAR v.2.1 - ERDAS IMAGINE 2013 - SPSS Statistics 20 - ArcGIS 10.1

Publicación

Montealegre, A.L., Lamelas, M.T., de la Riva, J., García-Martín, A., Escribano, F. 2015. Cartografía de modelos de combustible mediante combinación de imágenes LiDAR, SAR y ópticas en el Centro de adiestramiento "San Gregorio". In *Asorey Cacheda, R., Devesa Rey, R., Solla Carracelas, M.M., Pousada Carballo, J.M. (Eds), III Congreso Nacional de i+d en Defensa y Seguridad, DESEi+d 2015. Centro Universitario de la Defensa de Marín, Marín, pp. 1185-1192 (ISBN:978-84-944537-0-0).*

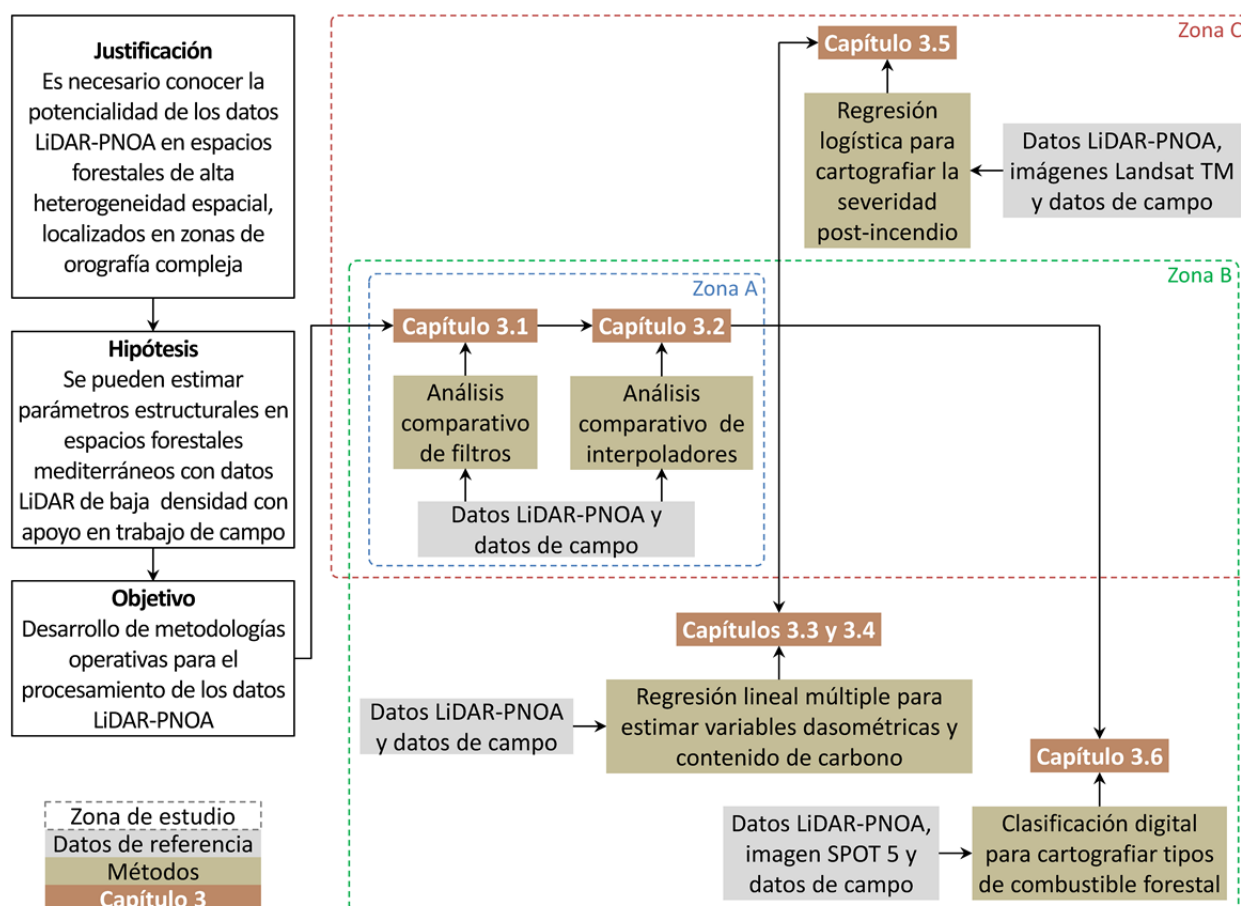


Figura 8. Esquema general de la estructura en el que se evidencia la interrelación entre los estudios realizados.

2 ÁREA DE ESTUDIO, MATERIAL Y MÉTODOS

Una vez se ha contextualizado la tecnología LiDAR, singularmente los sistemas aeroportados de pulsos discretos y huella pequeña, orientados al estudio de la estructura de la vegetación, y antes de dar cumplimiento a los objetivos específicos fijados en esta tesis doctoral, en el presente capítulo se describen, tanto el área de estudio, compuesta por tres zonas diferentes circunscritas a la Comunidad Autónoma de Aragón, como los materiales y las metodologías utilizadas en el trabajo de campo y de gabinete. Primeramente se detallan las características de los datos LiDAR del PNOA y a continuación se explica el funcionamiento de los algoritmos e interpoladores empleados en la investigación. Tras el tratamiento de los datos previo, se muestra cómo se ha llevado a cabo la validación de los resultados y cómo se pueden analizar las fuentes de error relacionadas con las propiedades de los datos y las peculiaridades del área de estudio. En los siguientes apartados se describe el trabajo de campo realizado para inventariar el bosque, valorar la severidad e identificar los tipos de combustible forestal. Posteriormente, se presentan las variables derivadas de las imágenes ópticas (índices espectrales) y de los datos LiDAR (métricas de los retornos láser) que se han empleado para llevar a cabo las aplicaciones más específicas. En este sentido, también se describen los métodos de regresión utilizados para relacionar el dato de campo con el derivado de las fuentes de teledetección. Por último, se detallan la representación cartográfica de las variables estimadas, así como las técnicas de clasificación digital para mapear los tipos de combustible.

2.1 Área de estudio

La Comunidad Autónoma de Aragón se sitúa en el noreste de Península Ibérica (ente los 39° y los 43° de latitud norte) y ocupa 47.720,3 km², de los que corresponden 15.636,2 km² a la provincia de Huesca, 17.274,5 km² a la de Zaragoza y 14.809,6 km² a la de Teruel. El total representa un 9,4% de la superficie de España, siendo así la cuarta comunidad autónoma en tamaño. Aragón limita por el norte con Francia (regiones de Mediodía-Pirineos y Aquitania), por el oeste con las comunidades autónomas de Castilla-La Mancha (provincias de Guadalajara y Cuenca), Castilla y León (provincia de Soria), La Rioja y Navarra, y por el este con las comunidades autónomas de Cataluña (provincias de Lérida y Tarragona) y Comunidad Valenciana (provincias de Castellón y Valencia).

La orografía de esta región tiene como eje transversal el valle del río Ebro, el cual discurre entre dos somontanos, el pirenaico al norte y el ibérico al sur, que conectan con las dos grandes formaciones montañosas de la Comunidad: el Pirineo y el Sistema Ibérico.

El Pirineo se dispone longitudinalmente en tres grandes unidades ordenadas de norte a sur: el Alto Pirineo (integrado por el Pirineo Axial y las Sierras Interiores), la Depresión Intrapirenaica y las Sierras Exteriores. El Pirineo Axial está compuesto por granitos, cuarcitas, pizarras y calizas, y contiene las

máximas altitudes de la cadena montañosa, entre las que destaca el Pico Aneto con 3.404 m. Desde las grandes cumbres surgen importantes valles fluviales en sentido perpendicular al eje montañoso atravesando el corredor de la Depresión Intrapirenaica, en cuyo reborde meridional destacan los relieves conglomeráticos de San Juan de la Peña (1.552 m) y Peña Oroel (1.769 m). Por último, las sierras exteriores están formadas por materiales predominantemente calcáreos, alcanzando alturas entre los 1.500 y los 2.000 metros (Peña and Lozano, 2004).

Las Bardenas, los Somontanos de Huesca y Barbastro y La Litera, enlazan la zona prepirenaica con el Valle del Ebro por su margen izquierda. La Depresión del Ebro, ocupada en épocas mesozoica y eocena por el mar, fue cubierta por materiales terciarios y sedimentos aluviales aportados por el río y los únicos relieves que rompen el paisaje llano son, por la margen izquierda, los Montes del Castellar y la Sierra de Alcubierre (800 m) y, por la derecha, los Altos de La Muela, La Plana y Sillero.

El Sistema Ibérico está constituido por un conjunto de macizos montañosos y altiplanicies que aíslan la Depresión del Ebro por su flanco sur y se escalonan hacia el centro de la Península. En Aragón se extiende desde la Sierra del Moncayo (2.314 m) por el oeste, hasta los Puertos de Beceite y las Sierras de Gúdar-Maestrazgo por el este, y las Sierras de Javalambre y Albarracín en el sureste. Las cuarcitas y pizarras paleozoicas son los materiales principales en los relieves de mayor entidad, mientras que el resto de estructuras están formadas por calizas y dolomías jurásicas y cretácicas. Esta configuración del relieve supone que el 60% de la provincia de Teruel tenga una altitud media de 900 m (Peña and Lozano 2004).

Aragón se enmarca dentro del ámbito del clima mediterráneo continental, caracterizado por inviernos fríos y veranos calurosos y secos, y una pluviometría irregular y escasa. Sin embargo, la orografía diversa del territorio perturba los valores propios del clima y da lugar a una variedad de ambientes climáticos que abarcan desde la alta montaña de los Pirineos al norte, con nieves permanentes en las cumbres más elevadas, hasta las zonas esteparias o semidesérticas, como los Monegros o los llanos de Belchite en la Depresión, pasando por el marcado clima continental de la zona de Teruel-Daroca. Tal como describe Cuadrat (2004), las características principales que definen el clima aragonés son:

- La aridez, resultado de una situación de cubeta encajada entre las cordilleras montañosas pirenaicas y del Sistema Ibérico, que produce un efecto de sombra pluviométrica en el sector de la Depresión del Ebro. Mientras en los Pirineos y Montes Universales se sobrepasan los 1.000 mm/m² de precipitación, en el sector central de la Depresión no se alcanzan los 300 mm/m².
- Los destacados contrastes de temperaturas, con estaciones extremas muy prolongadas (inviernos muy fríos y veranos calurosos), y de transición (primavera y otoño) cortas y variables como consecuencia del efecto de la continentalidad.
- La irregularidad de las lluvias, con alternancia de años secos y húmedos.
- Los vientos intensos y frecuentes que se encajonan en el valle medio del Ebro en dirección noroeste a sureste (cierzo) y en sentido sureste a noroeste (bochorno).

Por otro lado, Aragón destaca por una amplia representación de los diferentes ecosistemas presentes en el territorio nacional, debido a su mencionada diversidad topográfica, climática y litológica, así como por las transformaciones antrópicas del paisaje. Los bosques ocupan más de 1,5 millones de ha, de las que aproximadamente la mitad corresponden a coníferas, una sexta parte a frondosas y el resto son bosques mixtos. También se pueden encontrar bosques abiertos y matorrales que ocupan una superficie de 1 millón de ha.

En Aragón coexisten las regiones biogeográficas más importantes del reino Holártico: la región Eurosiberiana y la Mediterránea. En relación a esta última, comprende la Depresión del Ebro, los dos Somontanos, los valles fluviales de la margen derecha del Ebro, la Hoya de Teruel y las parameras

turolenses. En ella predominan las comunidades de matorral de medio y bajo porte, fundamentalmente romerales, junto a enebrales, y especies gipsófilas, nitrófilas y halófilas, ligadas a sistemas endorreicos y litologías margo-yesíferas y arcillosas. Las formaciones boscosas, están dominadas por masas de quercíneas, sabinars y pinares xéricos de *P. halepensis* y *P. nigra* (Longares 2004).

En términos generales, gran parte de los espacios forestales aragoneses se caracterizan por una considerable peligrosidad estructural ante el fuego, debida a la complejidad del relieve y a la vegetación xerófila asociada al clima mediterráneo de interior, en el cual destaca la falta de humedad ambiental, los períodos continuados de altas temperaturas con tormentas de fuerte aparato eléctrico y los vientos intensos en la época estival. También han influido el retroceso notable del pastoreo extensivo y del aprovechamiento de leñas ocurrido en las últimas décadas (López del Ríos and Bardají 2009).

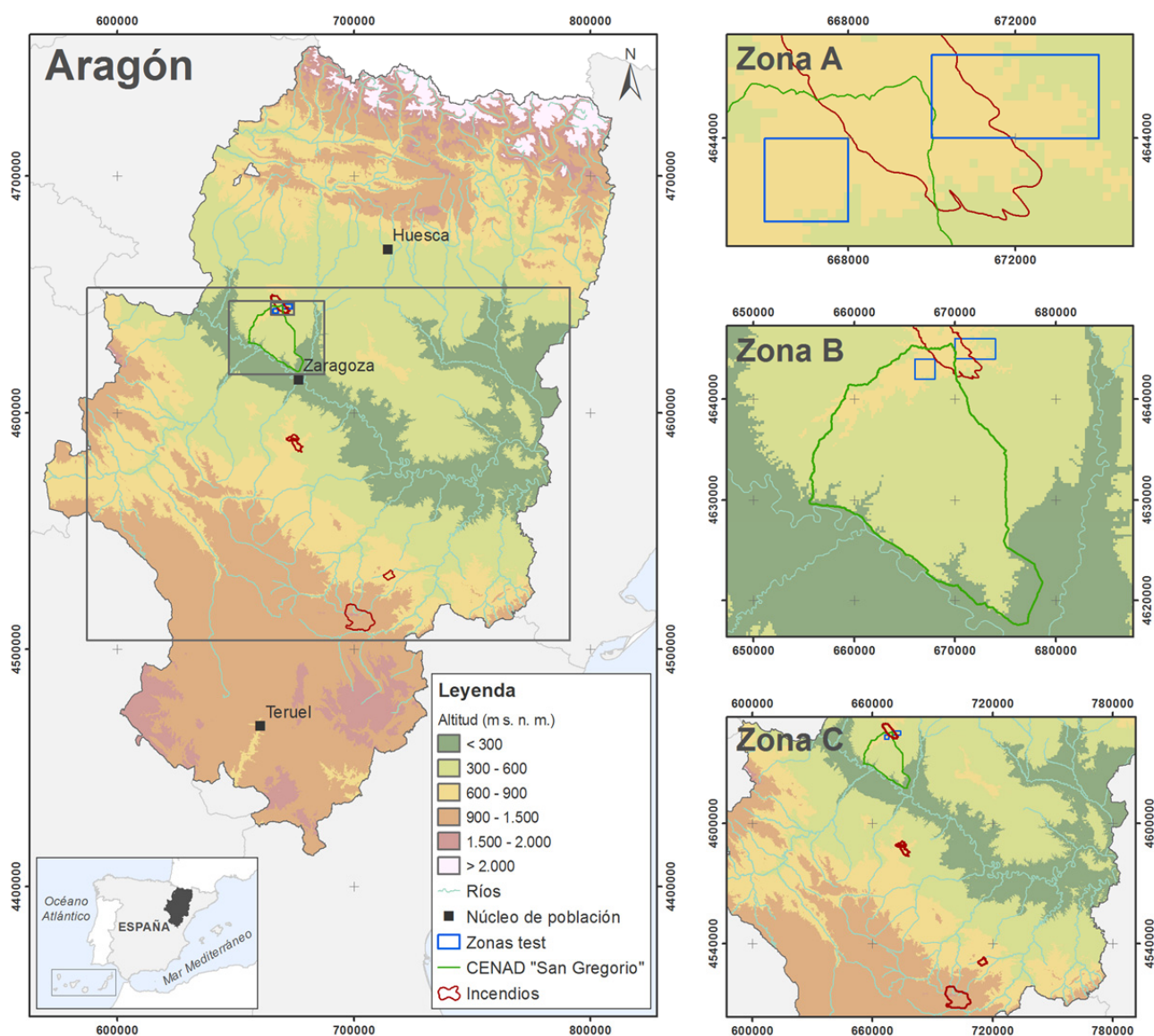


Figura 9. Localización general de las zonas de estudio A, B y C.

Tal como se ha comentado en el Capítulo 1, se han seleccionado tres zonas de estudio (ver Figura 9) interrelacionadas para dar respuesta a los objetivos específicos de la investigación, las cuales se describen a continuación:

- A. Zonas test, T1 (2x2 km) y T2 (4x2 km): situadas en el valle medio del Ebro, la región semiárida más septentrional de Europa, en el entorno de los Montes de Zuera (provincia de Zaragoza). El relieve de la zona presenta alturas entre los 400-750 metros sobre el nivel del mar y pendientes moderadas a fuertes. El paisaje es típicamente mediterráneo, donde los bosques monoespecíficos de *P. halepensis* Mill. cubren algo más de 700 ha, siendo el *Quercus coccifera* L. la principal especie del sotobosque. La vegetación arbustiva, dominada por esta última, se mezcla con *Juniperus oxycedrus* L. subsp. *macrocarpa* (Sibth. et Sm) Ball, *Thymus vulgaris* L. y con cultivos de cereales. Además, el sector occidental de T2 ha sido afectado por incendios forestales -en junio de 1995 y agosto de 2008-, estando actualmente cubierto por especies arbustivas rebrotadoras y restos de madera (ver descripción de la zona de estudio C para más detalles). A su vez, las zonas T1 y T2 están comprendidas dentro de la zona B, la cual se describe a continuación.
- B. Centro de Adiestramiento² (CENAD) “San Gregorio” y su entorno: se ubica en plena Depresión del Ebro, al norte de la ciudad de Zaragoza. Está concebido para apoyar a las Unidades de la Fuerza, especialmente las Acorazadas y Mecanizadas, en las diferentes fases de su preparación, para lo cual se despliegan todas las posibilidades de instrucción y adiestramiento, con simulación tanto real y constructiva, como virtual. Para cumplir con estas misiones, el CENAD “San Gregorio” cuenta con las instalaciones del Acuartelamiento “General Quintana” y el Área del Campo de Maniobras y Tiro de (ACMT) “San Gregorio”, que a su vez puede dividirse en tres zonas: Restringida (para uso prioritario de Unidades de la Plaza), Logística y de Maniobras, donde se encuentran las principales instalaciones de preparación, campos de tiro y zonas de caída de proyectiles.

El CENAD tiene su origen en el Campo de Tiro y Maniobras “Alfonso XIII”, creado en 1911, el cual fue ampliado a mediados de la década de 1970 hasta ocupar actualmente una superficie cercana a las 34.000 ha, que se integran casi en su totalidad en el Término Municipal de Zaragoza, y abarca cinco términos más: Zuera, Castejón de Valdejasa, Villanueva de Gállego, Tauste y Torres de Berrellén. El CENAD tiene forma trapezoidal, con una longitud mayor de 29 km y una anchura máxima de 22 km, aproximadamente. Está dominado al norte por los Montes de Castejón (760 m), a partir de los cuales se extienden glaciares, barrancos de diversa entidad³, vales y cabezos que acaban conformando un paisaje ondulado que se rompe por los escarpes sobre el cauce del Ebro. La cota más elevada del área de estudio se encuentra en el Vértice Esteban (742 m de altitud), mientras que los parajes más bajos (300 m de altitud) se localizan en el escarpe y en los cantiles que se asoman sobre los valles de los ríos Ebro (al sur) y Gállego (al este). La altitud media de la zona está comprendida entre 350 y 400 m. Cabe destacar que no existen cursos de agua permanentes y que los escasos acuíferos se recargan directamente a partir de las precipitaciones (Ministerio de Defensa 2007).

La mayor parte de la superficie del área de estudio es considerada forestal. Predominan los espartales, albardinales, lastonares, sisallares, ontinares y albardales, así como un matorral de tipo haloxerófilo y gipsícola. En menor medida se encuentran los pinares de *P. halepensis* en las zonas de los Montes de Zuera y Castejón de Valdejasa, así como en pequeños rodales

² Los Centros de Adiestramiento y los Campos de Maniobras y Tiro adscritos al Ministerio de Defensa, son las instalaciones que el Ejército de Tierra utiliza para la preparación de sus diferentes Unidades con la finalidad de llevar a cabo las misiones incluidas en los distintos planes operativos previstos, tanto actuales como futuros. En el caso concreto del CENAD “San Gregorio”, cabe destacar su especial relevancia debido a su reconocimiento como Zona de Interés para la Defensa Nacional (Real Decreto 538/2014, de 20 de junio), la más alta categoría en este ámbito, siendo un referente en la realización de maniobras y ejercicios como el *Trident Juncture* de la OTAN.

³ El barranco de la Virgen es un ejemplo paradigmático de incisión fluvial sobre yesos, que incluso afloran en forma de cristales, lo cual le ha valido su consideración como lugar de interés geológico (LIG) por la legislación autonómica del Gobierno de Aragón.

formando masas discontinuas. Junto a ellos, suelen encontrarse coscojares y otras especies termófilas como el *Juniperus*, que forman un matorral muy alto o maquia. Completan este paisaje vegetal heterogéneo la garriga, constituida por romerales, tomillares y aliagares de menos de 50 cm de altura (Ministerio de Defensa 2007).

En general, los centros de adiestramiento y los campos de maniobras y tiro son objeto de especial atención dada su naturaleza (el 96% de la superficie total de los terrenos de titularidad estatal adscrita al Ministerio de Defensa son forestales) y por la utilización de municiones y explosivos, lo que supone un riesgo directo para la ocurrencia de incendios (Sierra 2013). Por todo ello, en cumplimiento de la disposición adicional octava de la Ley 43/2003, de 21 de noviembre, de Montes, modificada por la Ley 21/2015, de 20 de julio de 2015, el Ministerio de Defensa cuenta con la Directiva 42/2010, de 12 de julio, de la Ministra de Defensa, que a su vez se aplica en el ámbito del Ejército de Tierra mediante la Directiva del JEME 06/12. Con esta normativa se pretende velar por la integridad del patrimonio natural, de las infraestructuras propias de este tipo de instalaciones y de las personas que desempeñan sus actividades en las mismas. Las directrices marcadas por el Ministerio de Defensa para conciliar las actividades de las Fuerzas Armadas con el respeto ambiental, se centran en la implantación de sistemas de gestión medioambiental sostenible para el cumplimiento de dicho objetivo.

En el caso concreto del CENAD “San Gregorio”, además del riesgo antrópico derivado de la propia actividad del ACMT (ejercicios de tiro, caída de proyectiles, acampadas, etc.), el riesgo de incendios forestales viene determinado por causas naturales (tormentas secas en época estival) y por negligencias y accidentes en la periferia del CENAD que pueden llegar a alcanzar el interior de las instalaciones militares. Por todo ello, el CENAD cuenta con un Plan Contra Incendios Forestales permanentemente actualizado y un Plan Técnico de Defensa Contra Incendios Forestales (aprobado por el Ministerio de Agricultura y Medio Ambiente el 30 de mayo de 2014).

- C. Cuatro incendios forestales de grandes dimensiones ocurridos entre los años 2008 y 2009; dos de ellos localizados en la provincia de Zaragoza (Zuera y Jaulín), y el resto en la provincia de Teruel (Aliaga y Los Olmos). En el incendio de Zuera, el fuego se inició tras un accidente de tráfico que se produjo en la carretera entre Villanueva de Gállego y Castejón de Valdejasa (Término Municipal de Zuera) y, como consecuencia, alrededor de 2.200 ha de bosque fueron calcinadas entre el 6 y el 16 de agosto de 2008. Cabe destacar que no era el primer siniestro que afectaba a esta zona, caracterizada por su gran recurrencia en las últimas décadas. De hecho, parte de la vegetación afectada correspondía al regenerado del gran incendio que se produjo en el año 1995, en el que ardieron más de 4.000 ha.

En el incendio de Jaulín, a unos 60 km en dirección sur del anterior, la causa del incendio fue la caída de un rayo a consecuencia de una tormenta de verano ocurrida el 21 de julio de 2009, quemándose alrededor de 1.800 ha.

El incendio de Aliaga se produjo el 22 de julio de 2009 tras una tormenta seca en la Sierra de Majalinos, en el Término Municipal de Aliaga. En aquel entonces, este incendio forestal fue el más extenso registrado en España en todo el año 2009, con 9.000 ha quemadas, siendo la peor catástrofe natural de Aragón desde el incendio de Villarluengo, en el Maestrazgo Turolense, en el que ardieron, en el año 1994, 18.000 ha.

Por último, el incendio de Los Olmos también responde a la misma causa que los dos anteriores. El 21 de julio de 2009 se inició el fuego en el Término Municipal de Los Olmos, el cual consumió 500 ha de terrenos forestales.

Ambos incendios de Zuera y Jaulín se sitúan en el sector central de la Depresión del Ebro, entre los 500 y los 750 m de altura sobre el nivel del mar, y presentan un ambiente mediterráneo continentalizado. La topografía ondulada o alomada es característica de este sector, alternándose los valles de fondo plano y los barrancos de incisión lineal con zonas elevadas más llanas. Los incendios destruyeron bosques de *P. halepensis* con presencia de sotobosque compuesto por *Quercus coccifera* L., *Juniperus oxycedrus* L., *Rosmarinus officinalis* L. y *Genista Scorpius*, así como pastizales y algunos campos de cultivo.

En el caso de los dos incendios de Aliaga y Los Olmos, se localizan en las estribaciones montañosas del Maestrazgo, en el centro-este de la provincia de Teruel, a una altitud que oscila entre los 1.600 y los 678 m de altura sobre el nivel del mar. En ambos casos, con un clima más continentalizado si cabe que en Zuera y en Jaulín, se observan grandes oscilaciones térmicas a lo largo del año, con unos veranos cálidos y secos e inviernos largos y fríos. Destacan las tormentas estivales con fuerte aparato eléctrico y las importantes nevadas que se registran todos los inviernos. En esta zona, el fuego consumió exuberantes bosques de *P. nigra*, *P. sylvestris* y *Quercus ilex*, además de rodales de *P. halepensis*, formaciones de *Juniperus phoenicea*, matorrales esclerófilos, pastizales y algunas tierras de labor.

Tras los incendios, los servicios provinciales y la Dirección General de Gestión Forestal del Departamento de Medio Ambiente del Gobierno de Aragón realizaron diversas tareas para restaurar determinadas áreas. En los casos en que los troncos eran aprovechables, se cortaron y se apilaron para ser arrastrados y cargados. Con los demás restos se procedió de tres maneras diferentes: se trituraron los árboles en pie mediante un cabezal desbrozador acoplado a una retroaraña, se astillaron los árboles previamente apeados con una astilladora y se construyeron fajinas o albarradas con los fustes y ramas en aquellas zonas con mayor pendiente topográfica (Jordán de Urriés and Ureta 2009).

2.2 Material y métodos

2.2.1 Datos LiDAR-PNOA

El Plan Nacional de Ortofotografía Aérea (PNOA) nació en el año 2004 con el objetivo de obtener ortofotografías aéreas digitales con resolución de 25 y 50 cm y MDE de todo el territorio español con un período de actualización de 2-3 años, según las zonas. Se trata de un proyecto cooperativo y cofinanciado entre la Administración General del Estado y las Comunidades Autónomas que se enmarca dentro del Plan Nacional de Observación del Territorio (PNOT). Este Plan Nacional engloba las actividades del Instituto Geográfico Nacional (IGN) en materia de fotogrametría, teledetección y ocupación del suelo. En la estructura del PNOT se pueden diferenciar las siguientes fases (Figura 10):

1. Obtención y tratamiento de imágenes aéreas, de satélite y MDE del territorio español, con resoluciones y ritmos de actualización optimizados desde un punto de vista económico, así como el desarrollo de las aplicaciones en las que se van a utilizar dichas imágenes. Estas coberturas se materializan en planes específicos, como el PNOA, que tiene como objetivo la obtención de MDE, ortofotografías y nubes de puntos LiDAR, y el Plan Nacional de Teledetección (PNT), destinado a obtener coberturas de imágenes de satélite del territorio nacional de alta, media y baja resolución.
2. Extracción y generación de información geográfica a partir de dichas imágenes y datos en función de la resolución espacial de partida, como la cartografía topográfica básica y derivada, el Sistema de Información sobre Ocupación del Suelo en España (SIOSE) y otros sistemas de

información temáticos.

3. Disseminación y difusión de toda esta información.

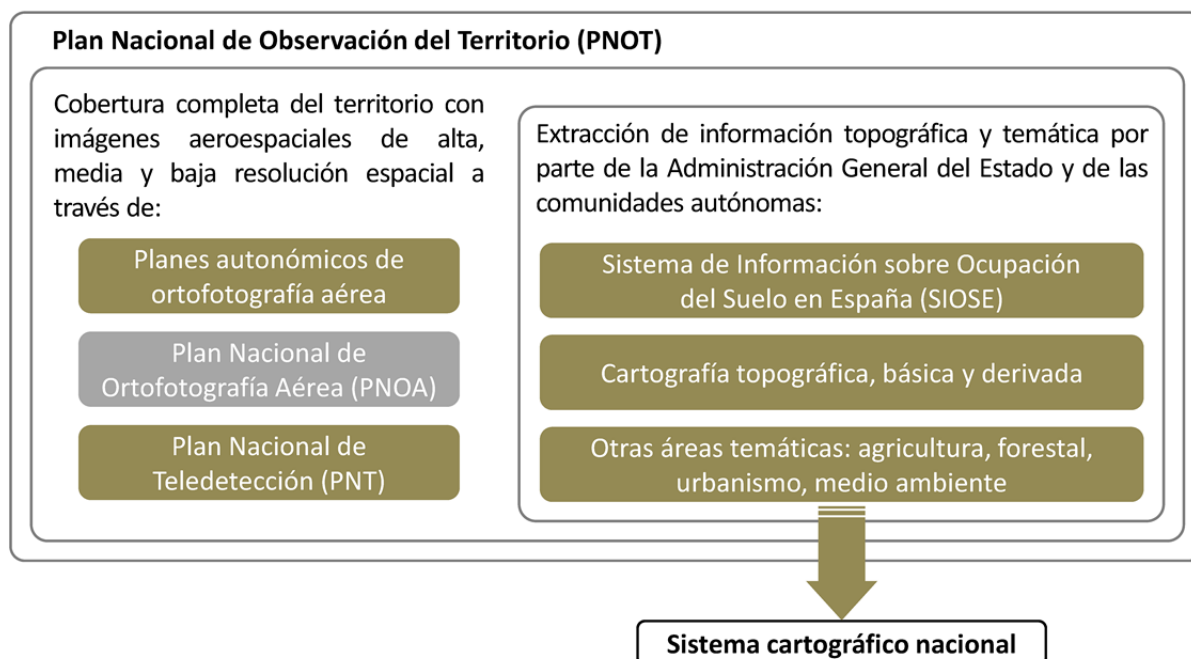


Figura 10. Contextualización del PNOA dentro del PNOT (CNIG 2017).

La dirección del PNOA es asumida por el Ministerio de Fomento, a través de la Dirección General del IGN y el Centro Nacional de Información Geográfica (CNIG), y se coordina con los demás Ministerios interesados y con cada Comunidad Autónoma, que a su vez coordina a las Consejerías competentes (Obras públicas, Agricultura, Medio Ambiente, etc.). El PNOA se basa en la realización de vuelos fotogramétricos y LiDAR únicos y en un tratamiento de los datos cumpliendo con unas especificaciones técnicas consensuadas entre todas las Administraciones Públicas participantes, para satisfacer las necesidades de todas ellas y evitar la duplicidad de esfuerzos y recursos. Por consiguiente, la producción se realiza de manera descentralizada por cada equipo autonómico, siendo el IGN el responsable de la validación final y de la integración de los productos resultantes. Este planteamiento de producción es acorde con los principios de la Directiva Inspire⁴ para el establecimiento de una Infraestructura de Información Espacial en la Comunidad Europea, que persigue que el nivel de detalle máximo de la información geográfica se capture una sola vez y que se comparta abiertamente entre los distintos agentes sociales. Tanto los recubrimientos de ortoimágenes como los MDE generados, forman parte de los “datos básicos de referencia” recogidos en el Anexo II de dicha Directiva [COM/2004/516/0175 (COD)]. La utilización de los mismos datos de partida por parte de las diferentes Administraciones permite obtener una perfecta coherencia geométrica y temporal de las bases de datos existentes, disponer de la información más actualizada y fomentar la

⁴ La Directiva Inspire (*Infrastructure for Spatial Information in Europe*), que entró en vigor el 25 de abril de 2007, establece las reglas generales para el establecimiento de una Infraestructura de Información Espacial en la Comunidad Europea basada en las Infraestructuras de los Estados miembros. Para asegurar que las infraestructuras de datos espaciales de los Estados miembros sean compatibles e interoperables en un contexto comunitario y transfronterizo, la Directiva exige que se adopten Normas de Ejecución comunes específicas para metadatos, conjuntos de datos, servicios de red, servicios de datos espaciales, datos y servicios de uso compartido y seguimiento e informes. Estas normas, implementadas a través de Guías Técnicas o Directrices, se consideran Decisiones o Reglamentos de la Comisión y por tanto son de obligado cumplimiento en cada uno de los países de la Unión (Geoportal IDEE 2017).

innovación y el desarrollo mediante la cooperación en proyectos entre Administraciones, universidades, organismos públicos de investigación y empresas.

En el caso concreto del vuelo LiDAR, hay que remontarse al año 2009, cuando se planteó la necesidad de obtener MDE de alta precisión mediante el uso de esta tecnología láser aeroportada para la realización de cartografía de áreas de inundación y proyectos de infraestructuras fundamentalmente (Figura 11). La cobertura LiDAR se ha realizado mediante colaboración y cofinanciación entre el Ministerio de Fomento (por medio del IGN y del CNIG), el Ministerio de Agricultura, Alimentación y Medio Ambiente (por medio de la Dirección General del Agua, las Confederaciones Hidrográficas y el Fondo Español de Garantía Agraria-FEGA) y el Ministerio de Hacienda y Administraciones Públicas (por medio de la Dirección General del Catastro), además de las Comunidades Autónomas.

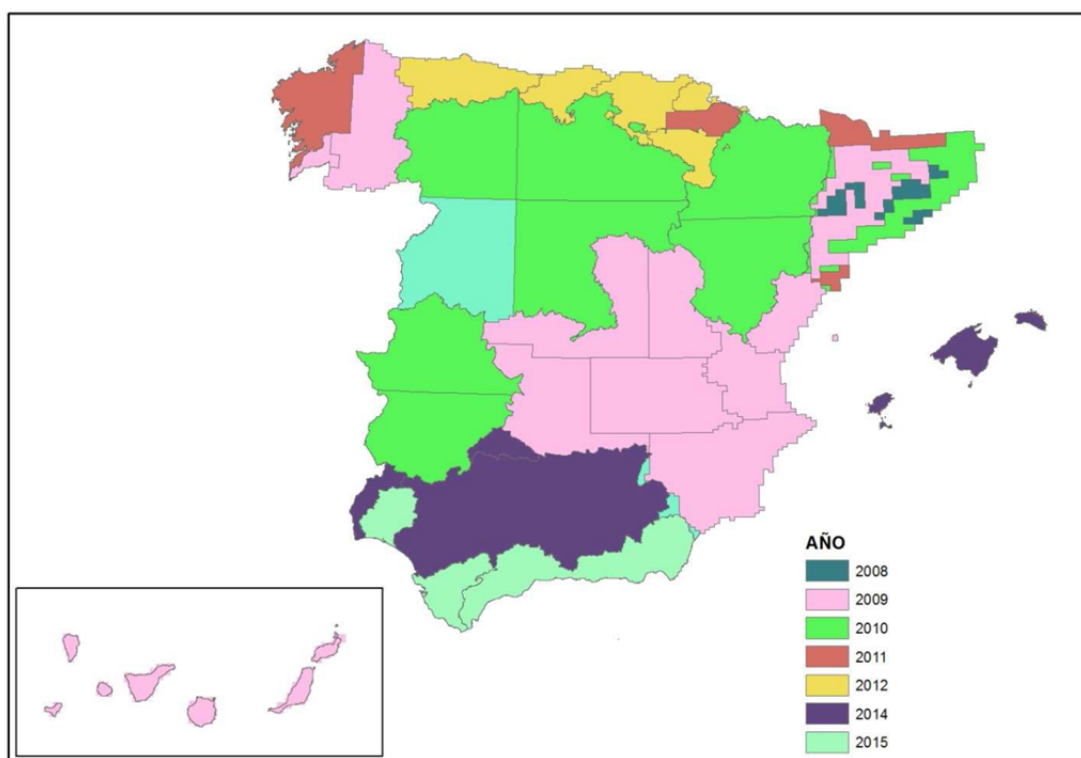


Figura 11. Cobertura LiDAR del proyecto PNOA 2008-2015 donde se muestran los años de captura de la información (CNIG 2017).

En un principio, sólo estuvieron accesibles a través del CNIG los MDE generados con LiDAR, pero posteriormente se pusieron a disposición de los usuarios las nubes de puntos de la práctica totalidad del territorio español, dado el interés y su potencial para desarrollar múltiples aplicaciones, tal como se ha descrito en el apartado introductorio de esta tesis. Está previsto continuar con la difusión de estos datos a medida que se procesen para las distintas zonas del territorio nacional (actualización del vuelo LiDAR cada 6 años dependiendo del presupuesto disponible). En este sentido, ya se ha iniciado la segunda campaña de vuelo y, en el caso concreto de Aragón, ya se ha volado la franja central del territorio de la Comunidad. Además, se pretende incrementar la densidad nominal de puntos, pasando de 0,5 puntos/m² (nubes de puntos actuales), a 1-2 puntos/m² en el futuro.

En la Tabla 3 se sintetizan las características técnicas básicas de las nubes de puntos del PNOA (CNIG 2017).

Tabla 3. Especificaciones técnicas del LiDAR-PNOA (CNIG 2017).

Característica	Descripción
Sensor	Leica ALS-50 y ALS-60; y RIEGL LMS-Q680.
Sistema geodésico de referencia	ETRS89 en la Península, Islas Baleares, Ceuta y Melilla, y REGCAN95 en las Islas Canarias.
Proyección cartográfica	UTM según el huso correspondiente a cada zona.
Modelo de geoide	EGM2008-REDNAP (adaptación del geoide mundial EGM08 a España) para realizar la transformación de cotas elipsoidales a ortométricas.
Campo de visión transversal (FOV)	El máximo FOV permitido es de 50° efectivos.
Frecuencia de escaneado	Mínimo de 70 Hz y hasta 40 Hz con un FOV de 50°.
Frecuencia de pulso	Mínima de 45 kHz, asumiendo un FOV de 50° y un máximo alcance de hasta 3.000 metros.
Densidad de puntos	0,5 puntos/m ² que implica un espaciamiento entre puntos $\leq 1,41$ m.
Resolución radiométrica de intensidades múltiples	Rango dinámico de al menos 8 bits.
Capacidad de detectar múltiples retornos para un mismo pulso	Hasta 4 retornos para cada pulso con una discriminación en distancia vertical de al menos 4 m.
Sistema de navegación basado en GPS	Equipo de GPS doble frecuencia de al menos 2 Hz.
Sistema inercial (IMU/INS)	Frecuencia de registro de datos ≥ 200 Hz y deriva $<0,1^\circ/\text{h}$.
Velocidad del avión en el momento de captura de los datos LIDAR	Variable.
Altura de vuelo	Hasta 3.000 m de alcance máximo.
Longitud máxima de pasada	90 km.
Precisión global horizontal y vertical nadiral después del procesado	La precisión global horizontal nadiral es inferior a 30 cm de RMSE _{x,y} (1 sigma) y la vertical nadiral es inferior a 20 cm de RMSE _z (1 sigma). En zonas de vegetación cerrada y pendientes acentuadas se admiten errores de hasta 3×RMSE. En los bordes del campo de visión se admiten precisiones del orden de 2×RMSE.
Precisión general altimétrica y error máximo	$\leq 0,40$ m en el 95% de los casos. No puede haber ningún punto con un error superior a 0,60 m.
Discrepancia altimétrica entre pasadas	$\leq 0,40$ m.
Formato de las nubes de puntos	Ficheros LAS ⁵ de acuerdo con teselas UTM de 2 km x 2 km de extensión. Clasificados automáticamente según el formato LAS, distinguiendo:
Clasificación automática de los ficheros LAS ⁶	- Suelo (terreno y puentes). - Vegetación. - Edificios. - Solape.

⁵ El formato de archivo LAS es un formato de archivo público para el intercambio de nubes de puntos entre los usuarios de datos. Este formato de archivo binario fue creado por la ASPRS como alternativa a los sistemas de propiedad o a los sistemas de intercambio de archivos ASCII genéricos utilizados por muchas empresas. Existen diferentes actualizaciones o versiones del formato LAS. La última especificación aprobada es la 1.4 (14 de noviembre de 2011) (ASPRS 2016).

⁶ En algunos casos, las nubes de puntos han sido coloreadas en RGB a partir de las ortofotografía del PNOA con tamaño de pixel de 25 o 50 cm.

Por lo que respecta a los atributos de los datos LiDAR (Figura 12), además del valor posicional x , y , z , se almacena información adicional asociada a cada punto. A continuación se describen los atributos de acuerdo con las especificaciones del formato LAS (ASPRS 2016):

- Intensidad: cantidad de energía reflejada por el retorno que da lugar a un punto y que se codifica con valores de 0 a 255.
- Número de retorno: un pulso láser emitido puede tener hasta cinco retornos según las características de la superficie y las capacidades del sensor. El primer retorno será marcado como retorno número 1, el segundo como retorno número 2, y así sucesivamente. Es decir, el número de retorno hace referencia a la numeración secuencial que identifica cada uno de los retornos registrados por el sensor.
- Número de retornos (para cada pulso emitido): número total de retornos por cada pulso dado. Por ejemplo, un punto puede tener un número de retorno 3 y corresponder a un pulso con un número total de cinco retornos, es decir, es el tercer eco de un total de cinco para un único pulso.
- Dirección de escaneo: sentido en el que se mueve el espejo del escáner láser en el momento de emisión del pulso. Un valor de 1 es una dirección de escaneo positiva, y un valor de 0 es una dirección de escaneo negativa. Por ejemplo, un valor positivo indica que el escáner se está moviendo del lado izquierdo al lado derecho de la dirección de vuelo.
- Borde de la línea de vuelo: identificación numérica para conocer si el punto está situado en el borde de la franja de escaneo, es decir, si es el último antes de que el escáner cambie de dirección (valor 1) o no (valor 0).
- Clasificación⁷: conjunto estándar de códigos predefinidos que permiten “etiquetar” cada uno de los puntos de la nube para identificar el tipo de objeto que ha producido la reflexión del pulso láser. Los códigos de clasificación han sido definidos por la ASPRS para los formatos LAS 1.1, 1.2, 1.3 y 1.4, y son los siguientes:
 - 0: Punto que nunca ha sido clasificado.
 - 1: Punto que, habiéndose sometido a un proceso de filtrado, no se ha asignado a ningún tipo de objeto.
 - 2: Suelo o terreno desnudo.
 - 3: Vegetación baja.
 - 4: Vegetación media.
 - 5: Vegetación alta.
 - 6: Edificaciones.
 - 7: Punto bajo (ruido).
 - 8: Punto clave que no debe retenerse en ningún algoritmo de simplificación.
 - 9: Agua.
 - 10: Reservado por la ASPRS.

⁷ Los datos LiDAR-PNOA proporcionados por el CNIG para el desarrollo de la investigación estaban sin clasificar. En el momento en el que se pusieron a disposición pública a través del sitio web del CNIG, los datos ya estaban clasificados. Fuentes consultadas señalan que se trata de un algoritmo de filtrado común para todo el territorio nacional basado en el filtro de Axelsson (2000) y que está implementado en el *software* propietario TerraScan®.

- 11: Reservado por la ASPRS.
- 12: Punto de solape.
- 13-31: Reservado por la ASPRS.
- Rango de ángulos de exploración: valor en grados entre -90 y +90. El valor 0 grados indica que el pulso láser está justo en el nadir. El signo negativo quiere decir que son pulsos situados a la izquierda del nadir, mientras que el signo positivo indica que el pulso láser está a la derecha del nadir.
- Datos del usuario.
- Identificador del origen del punto: valor que indica el archivo desde el que se originó el punto. Los valores válidos para este campo son de 1 a 65.535 inclusive.
- Tiempo GPS: fecha de registro del GPS en la que se emitió el pulso láser desde el avión. El tiempo está expresado en segundos.

Contents	Preview	Description												
	Sha	X Easting	Y Northing	Z Elevation	Intensity	ReturnNum	NumOfRetur	ScanDirFla	EdgeFlight	Classifica	ScanAngleR	UserData	PointSource	GPS TIME
► Point	664121,35	4642002,24	773,84	157	1	1	1	1	0	1	-11	138	17	-20270771,77
► Point	664003,55	4642110,66	748,38	107	1	1	1	0	0	1	-14	139	17	-20270770,15
► Point	664003,66	4642005,16	757,77	29	1	1	1	0	0	1	-11	137	17	-20270770,36
► Point	664094,88	4642086,1	775,9	155	1	1	1	0	0	1	-13	140	17	-20270771,28
► Point	664124,48	4642061,89	776,8	110	1	1	1	1	0	1	-12	138	17	-20270771,68
► Point	664013,27	4642111,33	755,47	38	1	2	1	1	0	1	-14	137	17	-20270770,27

Figura 12. Ejemplo de tabla de atributos de uno de los ficheros LAS empleados.

En total se ha trabajado con 314 archivos LAS proporcionados por el CNIG, cuyos principales parámetros se resumen en la Tabla 4. Un ejemplo de la nomenclatura que utilizan estos ficheros puede verse en la Figura 13. La época del año en la que fueron adquiridas las nubes de puntos no ha sido un inconveniente, dado que el ecosistema forestal estudiado está dominado por especies perennifolias.

Tabla 4. Principales características de los datos LiDAR-PNOA según el área de estudio.

Característica	Zona A	Zona B	Zona C
Fecha de adquisición (d/m/a)	22/01/2011 y 05/02/2011	23/01/2011 y 02/01/2011	<ul style="list-style-type: none"> • Zuera (10/01/2011; 02/02/2011; 05/02/2011) • Jaulín (01/08/2010) • Aliaga (27/12/2010; 24/01/2011; 21/01/2011) • Los Olmos (04/11/2010)
Sensor	Leica ALS-60	Leica ALS-60	Leica ALS-60 y ALS-50
Altura media de vuelo (m)	3012	3001	3355
Velocidad media del avión (km/h)	241	241	155
Ángulo de escaneo desde el nadir (grados)	±22	±29	±33
Superficie cubierta por los archivos LAS (km²)	12	1020	244

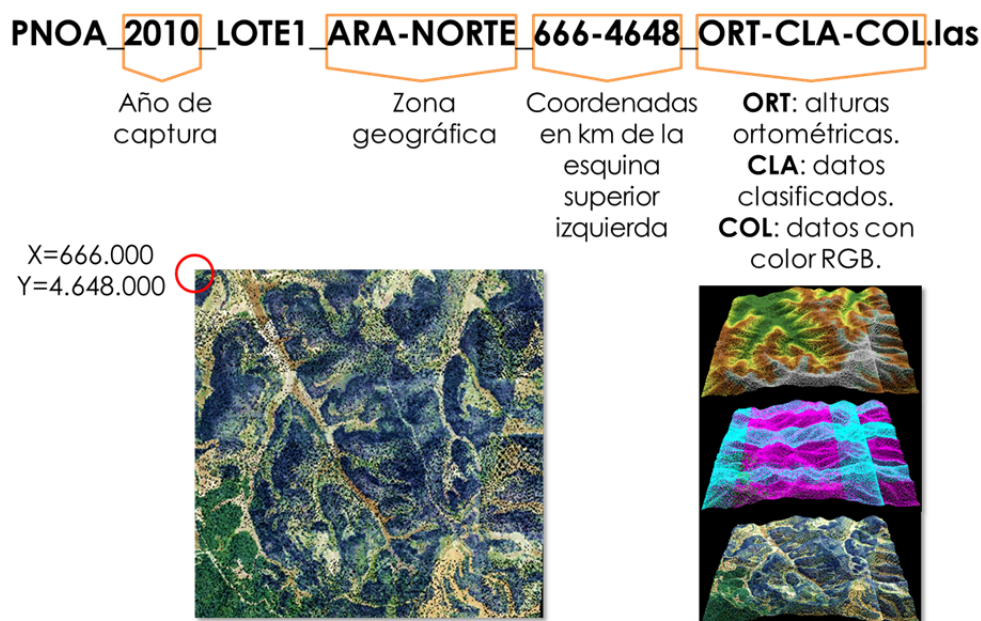


Figura 13. Ejemplo de fichero LAS del vuelo LiDAR del PNOA.

2.2.2 Algoritmos de filtrado

En este apartado se describen los algoritmos utilizados para filtrar las nubes de puntos LiDAR de la zona de estudio A. Nótese que se trata de filtros implementados en *software* libre a través de comandos o herramientas y que las parametrizaciones empleadas para su aplicación se describen en el Capítulo 3.

LAStools

Este conjunto de herramientas, que ha sido integrada en ArcGIS 10.1, contiene un comando llamado “*lasground*” que utiliza el algoritmo desarrollado por Axelsson (2000). En primer lugar, el algoritmo crea una matriz o malla regular (con una resolución determinada) que se superpone a la nube de puntos. Dentro de cada celda selecciona los puntos más bajos, los cuales son considerados “puntos semilla” o puntos iniciales de suelo, a partir de los cuales construir una primera superficie del terreno mediante una red de triángulos irregulares (TIN). Esta superficie es la que utilizará el algoritmo como referencia o primera superficie con la cual comparar el resto de puntos a clasificar. En cada iteración, y en cada triángulo, un punto no clasificado es añadido cada vez para redensificar el TIN (que se va construyendo con los puntos considerados de suelo), teniendo en cuenta que la distancia del punto a la faceta del triángulo, y el ángulo entre la faceta y la línea que conecta el punto con el vértice más próximo de la faceta, no deben exceder unos umbrales determinados (Figura 14, Vosselman and Maas 2010). En definitiva, antes de la siguiente iteración todos los puntos clasificados como suelo son añadidos al TIN, de manera que se genera un nuevo TIN más denso. Las iteraciones finalizan cuando todos los puntos quedan clasificados. La parametrización de *lasground* conlleva seleccionar la resolución de la matriz según el tipo de terreno, así como el refinamiento o nivel de exigencia en la búsqueda de los puntos iniciales.

En contraposición a los métodos pertenecientes al grupo de densificación progresiva de una superficie TIN, como el descrito anteriormente, donde la cantidad de puntos asignados a la clase de suelo se va incrementando paso a paso, los métodos basados en superficies interpoladas, suelen asumir que inicialmente todos los puntos de la nube pertenecen a la superficie de terreno para después, ser

iterativamente eliminados o reducida la influencia de aquellos que no se ajustan al modelo de superficie que paso a paso se va refinando en cada iteración. Los tres filtros que se explicarán a continuación son de este tipo: utilizan una superficie reconstruida a partir de la nube de puntos para llevar a cabo el proceso de filtrado de todos los puntos.

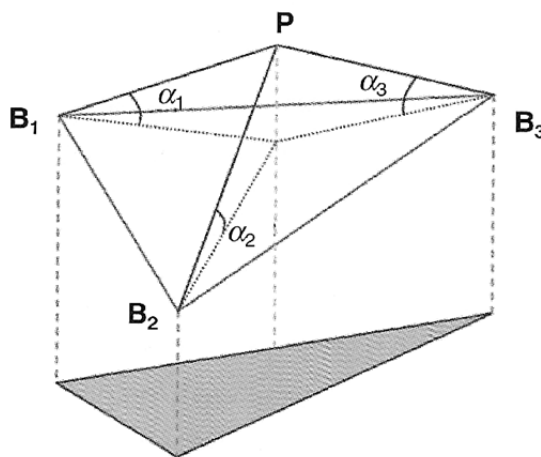


Figura 14. El punto P es evaluado utilizando los ángulos α_1 , α_2 , α_3 que se conforman con respecto al triángulo definido por los puntos B_1 , B_2 y B_3 , los cuales han sido clasificados como suelo y forman parte de la superficie TIN. Si el punto no superase los umbrales, sería clasificado como suelo y añadido al TIN. Tomada de Vosselman and Maas (2010).

FUSION v.3.30

El algoritmo que utiliza este *software*, implementado en el comando “*groundfilter*” y adaptado de Kraus and Pfeifer (1998), está basado en el cálculo de una superficie media usando todos los puntos de la nube, empleando una interpolación lineal robusta basada en el cálculo de mínimos cuadrados con un conjunto de funciones de pesos. En el primer paso, la superficie se calcula otorgando igual peso a todos los puntos, resultando una superficie topográfica irreal que se encuentra entre el verdadero suelo y la cubierta vegetal. Si los puntos quedan por encima de la superficie tendrán menos influencia sobre la forma de ésta que si los puntos quedan por debajo de dicha superficie, a los cuales, el filtro les dará más peso ya que tendrán residuales negativos. Paso a paso, se asignan pesos a los puntos, basados en la distancia y dirección que existe entre cada punto y la superficie calculada en la iteración anterior.

MCC-LiDAR v.2.1.

El algoritmo del cual toma el nombre este *software* (*Multiscale Curvature Classification*) fue desarrollado por Evans and Hudak (2007). Al igual que el filtro anterior, se fundamenta en un proceso iterativo basado en la definición de una superficie intermedia. El filtro iterativamente identifica los puntos de no suelo como aquellos que exceden umbrales positivos de curvatura a múltiples escalas.

Como se puede observar en la Figura 15, en primer lugar se crea una superficie inicial por interpolación empleando la elevación de los puntos de la nube $Z(s)$ (vector definido por las coordenadas x , y , z de cada punto LiDAR). Para ello se utiliza una función de base radial *thin-plate spline* (TPS-spline de lámina delgada), que es un tipo de interpolación exacta, es decir: la superficie debe pasar por cada punto. Hay dos parámetros que, en la fase inicial del filtrado, deben ser definidos por el usuario: la escala (λ) y el umbral o tolerancia de curvatura (t). El parámetro λ óptimo está relacionado con el tamaño de los objetos o elementos presentes en la escena, así como con el espaciamiento de

puntos de la nube. Esto es así porque este parámetro es el que define la resolución espacial de la superficie TPS interpolada. El valor de t , en cambio, requiere de cierta experimentación con diferentes valores entre 0,3 y 0,5.

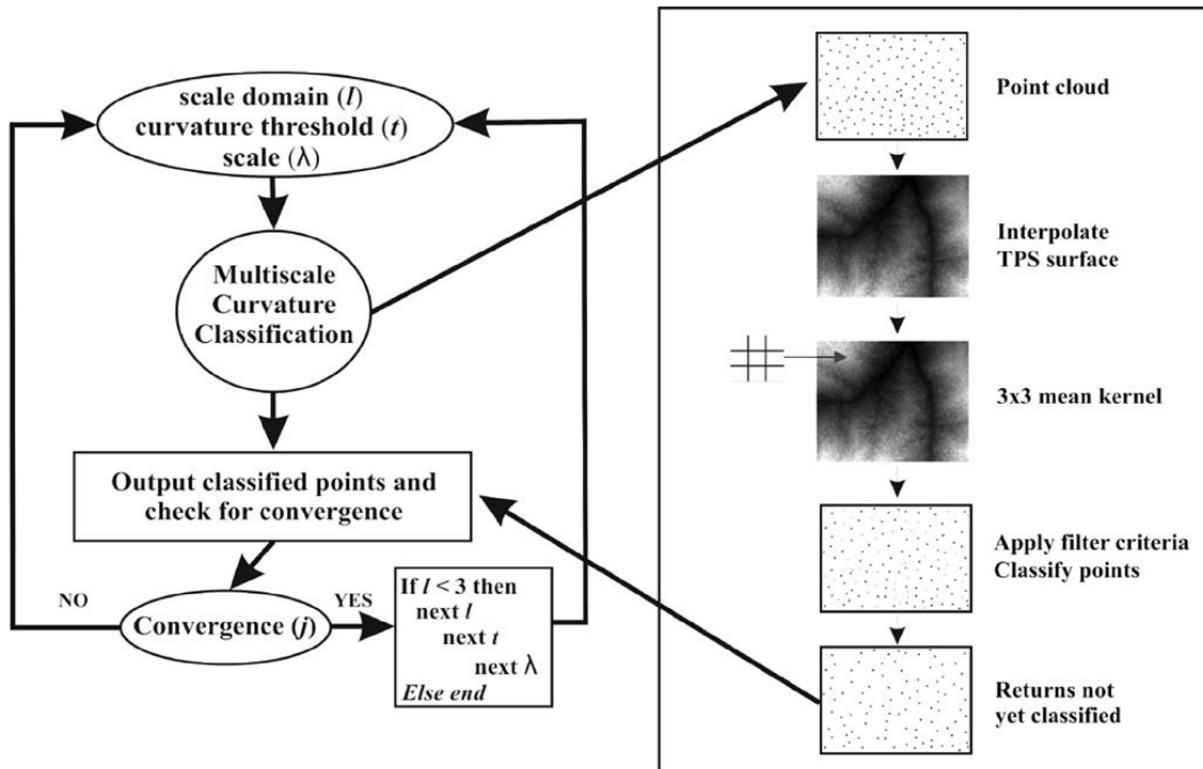


Figura 15. Diagrama de flujo del funcionamiento del algoritmo de filtrado MCC. Donde l es el dominio de escala, y t y λ son los parámetros que definen la tolerancia de curvatura y la escala, respectivamente. Tomado de Evans and Hudak (2007).

Después, un *kernel* u operador de media de 3×3 es pasado sobre la superficie interpolada y un nuevo vector $x(s)$ –definido por las coordenadas y el valor devuelto por el *kernel*– se crea de manera coincidente con $Z(s)$. En este paso se crea una nueva superficie a partir del cálculo de la curvatura de la superficie TPS celda por celda, ajustada a través de esa celda y sus vecinos circundantes. Entonces, si la altura de $Z(s) > x(s) + t$, donde $x(s)$ es el vector de la altura media coincidente con $Z(s)$ y t es la tolerancia de curvatura en el dominio de escala l , ese punto LiDAR $Z(s)$ es clasificado como no suelo y descartado del conjunto de puntos sin clasificar.

Posteriormente, el parámetro de convergencia j es evaluado para ver si el modelo debe seguir iterando o comenzar con el siguiente dominio de escala (l). El dominio de escala l es un bucle donde el conjunto de parámetros del algoritmo son redefinidos hasta converger. Así, el parámetro λ , que inicialmente debe definirse por el usuario, es recalculado en tres dominios de escala diferentes ($l=1$ a 3): primero como $0,5 \times \lambda$, segundo como λ y tercero como $1,5 \times \lambda$ (aproximación multiescala donde la superficie es interpolada a diferentes resoluciones). Al parámetro t le ocurre lo mismo: en el dominio de escala inicial ($l=1$) es definido por el usuario, mientras que en los dos siguientes dominios de escala es $t_1+0,1$ y $t_2+0,1$ (aproximación progresiva de tolerancia de curvatura). Una vez el modelo converge dentro de un dominio de escala, los parámetros λ y t son modificados, y el algoritmo procede con el siguiente dominio de escala hasta que el umbral de convergencia j es alcanzado en el último dominio de escala ($l=3$), que es cuando el filtro finaliza (ver Figura 15).

BCAL LiDAR Tools v.1.5.1.

Se trata de un conjunto de herramientas que incluyen un filtro iterativo relativamente sencillo, basado en superficies interpoladas, desarrollado por Streutker and Glenn (2006). Primero identifica los puntos con menor elevación a partir de un área de búsqueda definida por el usuario y después crea una superficie, mediante interpolación, con estos puntos. Dicho método de interpolación puede ser de tipo *splines*, distancia inversa ponderada, vecino natural y regresión polinómica. En cada iteración, cualquier punto que toca o queda debajo de la superficie interpolada en la anterior iteración es clasificado como suelo hasta que, paso a paso, ningún punto que cumpla esta condición quede sin clasificar. Al tratarse de un filtro iterativo, los puntos que se van agregando a la categoría de suelo van redefiniendo una nueva superficie que se interpola en cada iteración. Al final, los puntos que quedan por encima de la superficie son clasificados como no suelo.

La complejidad de esta herramienta de filtrado radica en la cantidad de parámetros que el usuario debe decidir: espaciamiento medio del dosel vegetal, nivel de refinamiento o cantidad de iteraciones, método de interpolación y umbral máximo de altura aceptable.

ALDPAT v.1.0.

Es un *software* que contiene diferentes algoritmos de filtrado que quedan englobados dentro de la categoría de filtros morfológicos, dado que se sirven de dos operaciones básicas de morfología matemática: “erosión” y “dilatación” aplicadas sobre una ventana móvil que se desplaza por la escena para clasificar los puntos de la nube. La erosión consiste en la selección del punto de cota mínima dentro de la ventana definida, mientras que la dilatación se basa en la selección del punto de cota máxima dentro de la ventana. Se pueden combinar estas operaciones para llevar a cabo una “apertura” (dilatación después de la erosión) y un “cierre” (erosión después de la dilatación) (Vosselman and Maas 2010).

- Filtro ETEW (*Elevation Threshold with Expand Window*): está basado en el algoritmo de Zhang and Whitman (2005). Primero la nube de puntos es subdividida en una matriz de celdas regulares y todos los puntos, excepto los de mínima elevación dentro de cada celda, son descartados. En la siguiente iteración, las celdas doblan su tamaño (esto se repite con cada iteración) y el filtro extrae una nueva elevación mínima dentro de ellas. Entonces, todos los puntos con elevaciones superiores a un umbral de altura sobre el mínimo son descartados. Es decir, para la iteración i^n , un punto $p_{i,j}$ es eliminado si $Z_{i,j} - Z_{i,min} > h_{i,T}$. Donde $Z_{i,j}$ representa la elevación del punto $p_{i,j}$ en una celda, $Z_{i,min}$ es la elevación mínima de esa celda y $h_{i,T}$ es el umbral de altura (relacionado con el tamaño de la celda en la iteración i^n y la máxima pendiente topográfica predefinida). Este proceso se repite con cada iteración, a medida que las celdas van duplicando su tamaño y van cambiando los umbrales de elevación, hasta que ningún punto de la iteración anterior es descartado. El usuario tiene que definir la resolución espacial de la matriz inicial, el factor de la pendiente y el número máximo de iteraciones.
- Filtro PM (*Progressive Morphological*): es un algoritmo desarrollado por Zhang *et al.* (2003) que se basa en el incremento gradual del tamaño de una ventana móvil de filtrado (que puede ser una línea unidimensional, un rectángulo bidimensional o cualquier otra forma) en combinación con el uso de umbrales de diferencia de elevación para eliminar los puntos que no pertenecen al terreno. En primer lugar, una malla rectangular es sobrepuesta a la nube de puntos. El filtro selecciona de entre todos los puntos de cada celda, el punto de menor cota p_j . Estos puntos constituyen una primera aproximación a la superficie del terreno. En segundo lugar, una operación de apertura es realizada sobre esa superficie para derivar una nueva superficie matricial. La diferencia de elevación ($dh_{i,j}$) de una celda j , entre las superficies previa ($i-1$) y última (i), es comparada con un umbral $dh_{i,T}$ para determinar si el punto candidato p_j en esa

celda no es de suelo. El umbral $d_{hi,T}$ es dependiente del error de las mediciones LiDAR, de la pendiente topográfica, del tamaño de la celda de la matriz y del tamaño de la ventana móvil de filtrado.

En el siguiente paso, el tamaño de la ventana de filtrado es incrementado y la superficie derivada en el paso anterior es utilizada como *input* para aplicar las operaciones de apertura. Este proceso se repite hasta que el tamaño de la ventana de filtrado es mayor que el tamaño máximo predefinido para los objetos o elementos presentes en la escena.

Como ocurre con este tipo de filtros implementados en ALDPAT v.1.0, el usuario debe definir el tamaño de celda y el de la ventana de filtrado, el factor de la pendiente y el umbral inicial de elevación.

- Filtro MLS (*Maximum Local Slope*): diseñado por Vosselman (2000), describe la diferencia máxima de altura admisible dentro de una ventana de filtrado circular, como una función de la distancia euclidiana entre dos puntos; el denominado concepto de pendiente local. Dado que la pendiente entre dos puntos de terreno es normalmente diferente de la pendiente observada entre el suelo y, por ejemplo, el ápice de los árboles, ésta pendiente o diferencia de gradiente puede ser utilizada para separar los puntos de suelo del resto. En definitiva, el filtro identifica los puntos de suelo comparando pendientes locales entre un punto candidato y sus vecinos. El filtro genera una matriz en la que cada punto de la nube es asignado a una celda. Si hay más de un punto dentro de la misma celda, el que tenga la menor cota será seleccionado. El punto es clasificado como suelo si el valor máximo de la pendiente entre este punto y otro punto (diferencia de altura) dentro de un radio determinado (distancia), es inferior a un umbral determinado.

2.2.3 Métodos de interpolación espacial

El LiDAR hace un submuestro aleatorio de la superficie que explora siguiendo un patrón de escaneo determinado, de ahí que sean necesarias dos tareas fundamentales para generar un MDE: separar o filtrar los retornos del suelo desnudo o terreno del resto de información (vegetación, elementos antrópicos) e interpolar dichos datos para producir una superficie continua, habitualmente en formato ráster (Vosselman and Maas 2010).

Tal como se comentó en el Capítulo 1, estudios previos han demostrado que la precisión de los MDE derivados del LiDAR no solo está influenciada por los procesos que implican la manipulación de los puntos (método de filtrado y de interpolación elegidos), sino también por una serie de factores externos, tales como la variabilidad topográfica, la densidad de retornos de suelo, la complejidad estructural del dosel vegetal y la resolución espacial. Todo ello afecta a la magnitud del error en la interpolación y da lugar a MDE con calidades diferentes en términos de fiabilidad con respecto a la realidad que describen (Bater and Coops 2009). Por estas razones, conocer la calidad interna y externa de un MDE generado con datos LiDAR es esencial, especialmente en los espacios forestales donde es necesario para normalizar las alturas de la vegetación.

Teniendo en cuenta estas consideraciones, se han examinado las seis rutinas de interpolación más habituales cuando se trabaja con datos LiDAR en el área de estudio A (zonas test T1 y T2), para poder identificar el interpolador y la resolución espacial más adecuados y conocer la influencia de varios factores en la precisión del MDE, dando así cumplimiento al objetivo 2 de la tesis doctoral, y sirviendo de apoyo al siguiente objetivo relacionado con la estimación de variables dasométricas en ese mismo ambiente mediterráneo.

Conviene señalar que los puntos de suelo desnudo utilizados fueron los resultantes de aplicar el algoritmo implementado en el *software* MCC-LIDAR 2.1, dado que fue el más idóneo para ambientes forestales y de topografía compleja y es el que proporcionó mejores resultados. De todos los puntos de suelo disponibles, se extrajo un 80% aleatoriamente para utilizarlos a modo de muestra de entrenamiento para generar los MDE. Teniendo en cuenta la densidad de puntos de las zonas test, se decidió generar MDE en formato ráster con una resolución espacial de 1 y 2 m.

Los métodos de interpolación utilizados en este estudio han sido los que se describen a continuación, de los cuales, los cinco primeros están implementados en ArcGIS 10.1 y el último en el *software* para aplicaciones forestales FUSION 3.30:

1. El vecino natural se basa en encontrar el subconjunto de retornos de suelo más cercano a un punto desconocido u objetivo aplicando ponderaciones sobre éstos mediante áreas proporcionales para interpolar un valor (Sibson 1981). Los vecinos naturales de un punto son aquellos que están asociados con los polígonos de Voronoi (Thiessen) adyacentes. Al principio, el algoritmo construye un diagrama de Voronoi de todos los puntos dados, y después crea un nuevo polígono de Voronoi alrededor del punto objetivo. La proporción de superficie de este nuevo polígono que queda superpuesta a los polígonos iniciales es la que se utiliza para asignar los pesos. De este modo, no es necesaria ningún tipo de parametrización. Además, es un método de interpolación local porque no infiere tendencias ni produce picos, depresiones, crestas o valles que no estén ya representados por los valores de altura de los retornos de entrada. En definitiva, el MDE que se genera “toca” a todos los retornos de suelo de entrada (Sambridge *et al.* 1995).
2. El TIN a ráster se basa en dos procedimientos. Primero se representa la superficie del terreno con triángulos irregulares continuos uniendo los retornos (ahora entendidos como nodos o vértices) mediante aristas, siguiendo el procedimiento de triangulación de Delaunay. Cada triángulo está vacío, por lo que no contiene ninguno de los puntos muestreados. En segundo lugar, se transforma el TIN en un ráster mediante dos alternativas posibles: bien se calculan los valores de celda al aplicar una interpolación lineal a los triángulos del TIN, o bien se calculan los valores de celda con una interpolación de vecinos naturales de los triángulos del TIN. Para esta investigación se testaron ambos procedimientos. Aunque la estructura matricial no puede representar puntos singulares ni estructuras lineales como el TIN, el proceso aprovecha parcialmente la capacidad del TIN para integrar discontinuidades en la generación del MDE, por lo que es preferible a la generación directa de la matriz regular a partir de los datos de entrada (Renslow 2013).
3. El método de ponderación en función inversa de la distancia (IDW) determina el valor de la celda a través de una combinación ponderada linealmente de un conjunto de puntos de entrada dentro de un radio de búsqueda determinado. La ponderación es una función de la distancia inversa elevada a una potencia matemática. El parámetro “potencia” o *power*, controla cómo la influencia de la ponderación disminuye mientras la distancia hacia el punto desconocido se incrementa. Al definir un valor de potencia alto, se puede poner el énfasis en los puntos más cercanos, por lo que tendrán más influencia sobre el punto desconocido y en general el MDE tendrá más detalle (la superficie será menos suave). Al especificar un valor bajo de potencia, los puntos circundantes presentarán más influencia, lo que resultará en una superficie más suavizada. Aunque el coeficiente de ponderación predeterminado es 2, el IDW fue aplicado con una potencia de 0,5 a 3 en intervalos de 0,5 para ponderar las distancias, teniendo en cuenta que el valor óptimo para la potencia es cuando el error medio absoluto mínimo es el más bajo. Las características de la superficie interpolada también pueden controlarse limitando los puntos de entrada que se utilizan en el cálculo de cada valor de celda

de salida. Esto puede ayudar a aumentar la velocidad de procesamiento, especialmente cuando se manejan muestras con millones de puntos LiDAR. Concretamente se especificó un radio variable dentro del cual se incluirían 6, 12 o 24 puntos para calcular el valor de la celda interpolada, con el fin de sopesar la variabilidad espacial de la muestra (Guo *et al.* 2010, Renslow 2013).

4. El ANUDEM o Topo a ráster es un método de interpolación diseñado específicamente para crear MDE hidrológicamente correctos. Está basado en el programa ANUDEM desarrollado por Hutchinson (1989, 2011). Este método utiliza una técnica de interpolación de diferencia finita iterativa. Se optimiza para tener la eficacia computacional de los métodos de interpolación local, como IDW, sin perder la continuidad de la superficie de los métodos de interpolación global, como *kriging* y *spline*. La opción de aplicación de drenaje se puede establecer para quitar todos los sumideros o depresiones de manera que pueda crearse un MDE hidrológicamente correcto. En este caso, se realizó el MDE aplicando, tanto esta opción, para que el algoritmo intente quitar todos los sumideros que encuentra, ya sean reales o espurios, como la opción que ignora el proceso de eliminación de sumideros.
5. El *kriging* ordinario, es un método de interpolación geoestadístico que asume que la distancia o la dirección entre los puntos de la muestra reflejan una correlación espacial que puede servir para explicar la variabilidad de la superficie a interpolar. El algoritmo ajusta una función matemática a una cantidad determinada de puntos o a todos los puntos dentro de un radio específico para determinar el valor de salida para cada celda. Es un método complejo, que conlleva varios pasos: el análisis estadístico exploratorio de los datos, el modelado de variogramas, la creación de la superficie y el análisis de su varianza.

El método *kriging* es similar al IDW, por cuanto pondera los valores de los puntos circundantes a una ubicación sin datos. Sin embargo, difiere en que las ponderaciones están basadas, no solo en la distancia entre los puntos medidos y la ubicación de la predicción, sino también en la disposición espacial general de los puntos. Para utilizar la disposición espacial en las ponderaciones, el método calcula la correlación espacial.

Básicamente, para llevar a cabo una predicción, el método crea los variogramas y las funciones de covarianza para calcular los valores de dependencia estadística (autocorrelación espacial), que dependen del modelo de autocorrelación, el cual debe ajustarse mediante una función. En este caso se utilizó una función normal o gaussiana para definir la variación espacial en el área en estudio, sectores de 1, 4, 4 con un *offset* de 45°, y 8, con 2 a 5 vecinos (Cressie 1988, Caruso and Quarta 1998).

6. “Punto a ráster” crea una superficie en la que cada celda representa la elevación media de los puntos que quedan contenidos dentro de la misma. Es la manera más simple de convertir los puntos a una superficie continua, pero se corre el riesgo de que existan celdas sin valor asignado si no hay, al menos, un retorno láser dentro de ellas (McGaughey 2009, Renslow 2013).

Se puede afirmar que no existe un algoritmo de interpolación que *a priori* sea claramente superior a los demás, puesto que cada uno posee ventajas y desventajas propias, y la elección del método por parte del usuario depende fundamentalmente del tipo de dato, del grado de precisión deseada (superficie más o menos suavizada), y de la cantidad de esfuerzo computacional que se puede asumir. De ahí el interés de esta investigación por facilitar la elección del más adecuado para la generación de un MDE lo más “realista” posible con respecto a la superficie original.

2.2.4 Validación de los métodos de filtrado e interpolación y análisis del error

El proceso de filtrado de una nube de puntos LiDAR puede dar lugar a errores, por lo que conviene examinar los resultados siguiendo un procedimiento metodológico exhaustivo (Sithole and Vosselman 2004), como el que proponen Zhang and Whitman (2005) y Zhang *et al.* (2003) a partir de la selección aleatoria de una muestra de puntos de la nube antes de ser filtrada para clasificarlos manualmente.

Un conjunto de 50 localizaciones fueron seleccionadas dentro de los límites de las áreas test T1 y T2, a partir de las cuales se trazaron unos *buffer* de 3 m de diámetro para extraer un total de 424 puntos LiDAR. Estos puntos fueron clasificados utilizando una combinación de fuentes de información: la ortofotografía del PNOA correspondiente (hoja 0322) del año 2009 de 0,5 m de paso de malla, los valores de intensidad y los valores de altura normalizada (ver Figura 16).

Aquellos puntos dificultosos, que no pudieron identificarse como terreno (clase 2 según el formato LAS) o no asignado (clase 1), tuvieron que verificarse mediante el replanteo de los mismos en jornadas de trabajo de campo empleando un equipo de posicionamiento global (GNSS-*Global Navigation Satellite System*) de la marca Leica Geosystems®, compuesto por un receptor VIVA GS15 y un controlador CS10 (Figura 17). Se trata de un sistema GNSS de altas prestaciones capaz de tomar mediciones de código y fase independientes en todas las frecuencias GPS y GLONASS. Por ejemplo, la precisión en términos de error medio cuadrático del GPS diferencial basado en la transmisión de diferencia de fase en tiempo real (RTK) en el modo estático rápido, es de 5 mm en horizontal y de 10 mm en vertical. Se utilizó el servicio de corrección de la Red Geodésica Activa de Aragón (ARAGEA) para conseguir un posicionamiento preciso en el mismo instante en que se efectuaba el trabajo de campo, dado que el receptor GNSS GS15 admite correcciones en el estándar RTCM (versiones 2.3 y 3.0) o CMR+ mediante una conexión a Internet (tarjeta de datos 3G) y cuenta con la funcionalidad NTRIP (*Networked Transport of RTCM via Internet Protocol*). Evidentemente, la precisión y fiabilidad de las mediciones tomadas en campo han dependido de varios factores, como el número de satélites disponibles, las obstrucciones, el tiempo de observación, la precisión de las efemérides, las condiciones ionosféricas, etc.

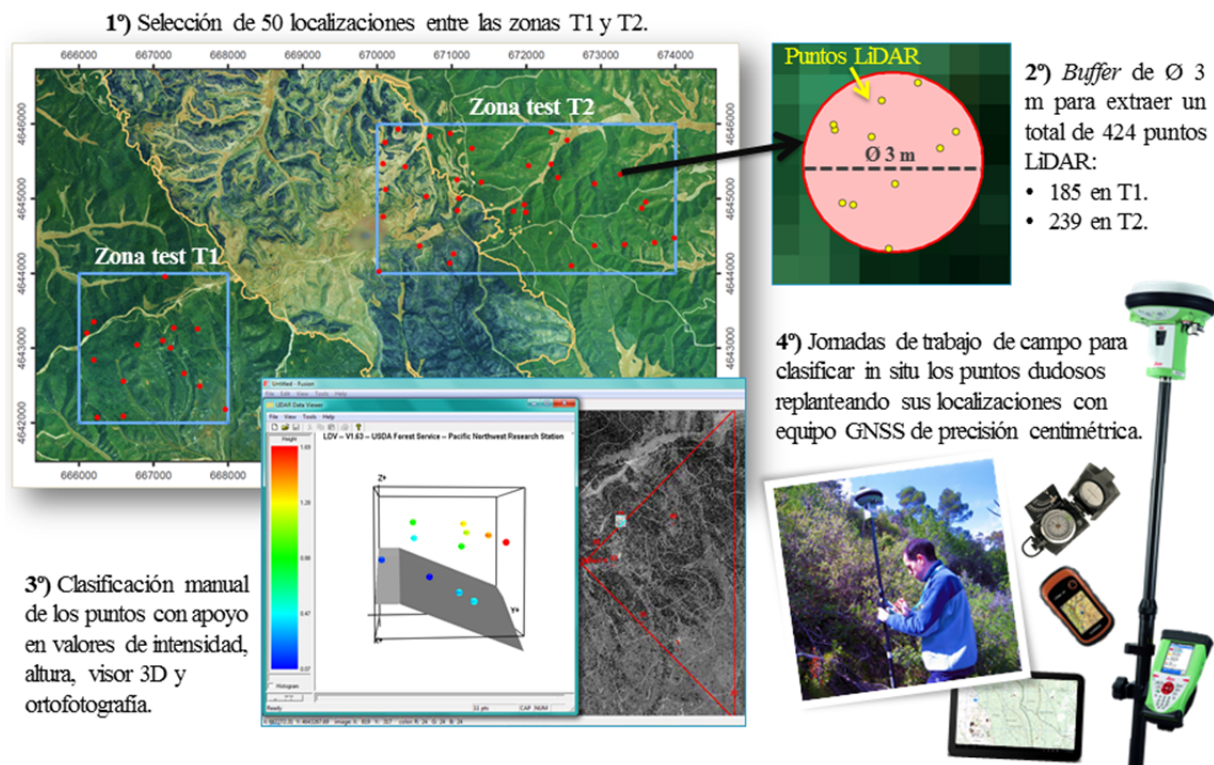


Figura 16. Etapas del proceso metodológico llevado a cabo para clasificar manualmente los puntos LiDAR.



Figura 17. Equipo GNSS de Leica Geosystems® compuesto por un receptor VIVA GS15 y un controlador CS10.

Conocer la fiabilidad global del filtrado es importante para saber si el método aplicado es el más adecuado (ecuación 4); sin embargo, puede enmascarar importantes diferencias entre categorías (suelo y no suelo), que escondan niveles de exactitud muy diversos. Por ello, un análisis más riguroso debe también considerar el número de puntos que, perteneciendo a una determinada categoría, no fueron incluidos en ella (Figura 18). El error de omisión o error Tipo I (ecuación 5) es aquel que permite cuantificar la imprecisión cometida por el filtro al no haber incluido los puntos que son realmente de suelo en dicha categoría. Por otro lado, también se puede calcular el error de comisión o error Tipo II; esto es, puntos que se incluyeron en una determinada categoría perteneciendo realmente a otra (ecuación 6). Si éste es elevado quiere decir que el filtro ha realizado una clasificación excesivamente amplia. En definitiva, ambos tipos de error expresan dos enfoques complementarios sobre el error en el filtrado de la nube de puntos (Sithole and Vosselman 2004).

$$Error\ Total = \frac{a + b}{N_{suelo} + N_{no\ suelo}} \quad (4)$$

$$Error\ Tipo\ I = \frac{a}{N_{suelo}} \quad (5)$$

$$Error\ Tipo\ II = \frac{b}{N_{no\ suelo}} \quad (6)$$

Donde a es el número de puntos de suelo incorrectamente clasificados como no suelo, b es el número de puntos de no suelo erróneamente clasificados como suelo, N_{suelo} es el número total de puntos de suelo en la muestra de referencia y $N_{no\ suelo}$ es el número total de puntos de no suelo en la muestra de referencia.

Por otro lado, el índice Kappa es una buena técnica para evaluar si el filtro ha discriminado los puntos de suelo y no suelo con una exactitud significativamente mayor a la que se hubiera obtenido con una asignación aleatoria. Un valor de Kappa igual a 1 indica un acuerdo pleno entre la realidad y el punto clasificado, mientras un valor cercano a 0 sugiere que el acuerdo observado es puramente debido al azar (Pardo and Ruíz 2005).

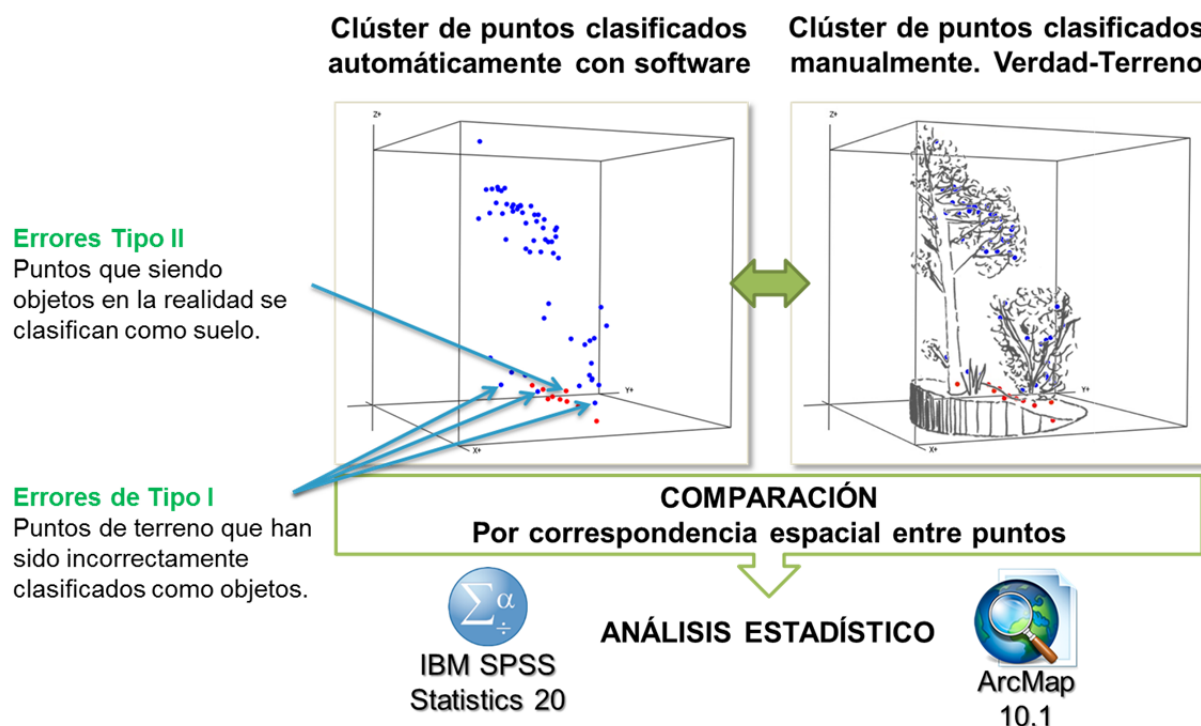


Figura 18. Ejemplo del análisis estadístico sobre el error en el filtrado. Como se puede apreciar, se realizó una comparación entre los puntos clasificados por el filtro y los clasificados manualmente mediante un análisis de correspondencia espacial. A partir de ahí, se trató la información en el *software* de estadística SPSS Statistics 20 para calcular los estadísticos del error.

Si estos estadísticos, que miden el grado de acuerdo entre categorías, se aplican de una manera estratificada considerando los factores más importantes que influyen en el filtrado -pendiente topográfica, tipo de cubierta forestal, densidad de puntos, y ángulo de escaneo-, el análisis del error se torna mucho más completo.

La calidad de un MDE depende del tipo y magnitud de los errores que inevitablemente se producen, debidos a que son visiones simplificadas de la realidad sometidas a un proceso de generalización. Los errores en los MDE pueden separarse en dos categorías: posicionales, que implican una deficiente localización geográfica de la cota y afectan a la situación en el plano x, y ; los atributivos, que suponen una asignación imprecisa de la altitud asociada a la cota e implican a las coordenadas en el eje z (Márquez 2004).

Ambos tipos de errores aparecen en el proceso de creación de los MDE, por lo que es necesario adoptar criterios para su detección y medida, así como para valorar la influencia de los mismos en los modelos derivados (propagación del error). Los errores posicionales afectan exclusivamente a los MDE vectoriales, por ejemplo TIN, mientras que en los matriciales, las localizaciones están predefinidas por su naturaleza ráster. Los errores atributivos pueden encontrarse en ambos tipos de MDE, pero son más habituales en los ráster, dado que la altura es asignada mediante operaciones de interpolación cuyo error es dependiente de múltiples factores (Márquez 2004).

La cuantificación del error sobre un MDE matricial puede realizarse por comparación de un conjunto de valores conocidos de altitudes con los correspondientes puntos homólogos del modelo, usando pruebas estadísticas convencionales (Tabla 5). En este sentido, el error de altitud en un punto $E(x,y)$ o residuo, se define como la diferencia entre ambos datos: $P_z(x,y) - M_z(x,y)$, donde el primer dato es la altitud del punto predicho por el MDE y el segundo es la altitud "real" o de referencia, que puede ser tomada de una fuente de información más precisa, como por ejemplo la medida directamente en el

campo, o bien extraerse de la misma fuente de información empleada para construir el modelo, siempre y cuando no haya intervenido en la generación del MDE.

En consecuencia, fue analizada la calidad interna y externa de los MDE generados tras el filtrado de las nubes de puntos. En el caso de esta última, se llevó a cabo un trabajo de campo utilizando el equipo GNSS de alta precisión descrito anteriormente. En total, se visitaron 55 localizaciones distribuidas aleatoriamente entre las zonas test T1 y T2 con el fin de tener una muestra de puntos de control con los cuales poder comparar los valores de altura predichos por el MDE y los medidos en la realidad (Aguilar and Mills 2008, Liu 2011).

Tabla 5. Estadísticos de error empleados para la evaluación objetiva de la calidad de los MDE.

Estadístico	Descripción	Utilidad
Error medio (EM)	$EM = \frac{1}{n} \sum_{i=1}^n (P_z(x, y) - M_z(x, y))$	<p>Donde n es el número de puntos que se comparan, $P_z(x, y)$ es la altitud del punto predicho y $M_z(x, y)$ es la altitud medida o de referencia del punto.</p> <p>El uso del EM sólo tiene interés para comprobar si las desviaciones del modelo son aleatorias o sistemáticas porque esta medida descriptiva, que tiene en cuenta el signo de los residuos, tiende a cero si hay magnitudes similares de valores positivos y negativos (error aleatorio). En caso contrario, si el EM es diferente de cero, puede tratarse de un error sistemático.</p>

Tabla 5. Continuación.

Estadístico	Descripción	Utilidad
Error medio absoluto (EMA)	$EMA = \frac{1}{n} \sum_{i=1}^n (P_z(x, y) - M_z(x, y))$	<p>La expresión matemática del EM se puede transformar tomando los valores absolutos de las desviaciones. Al no considerarse el signo, el valor del EMA siempre es mayor a cero y refleja el rango o distribución de los residuos. El 50% de éstos se localizará en el rango de ± 1 EMA.</p>
Error cuadrático medio (ECM)	$ECM = \sqrt{\frac{1}{n} \sum_{i=1}^n (P_z(x, y) - M_z(x, y))^2}$	<p>En una analogía con la desviación estándar, se obtiene el error de la raíz cuadrada de la media o la desviación de la raíz cuadrada media.</p>

Por otro lado, la existencia de una orografía compleja, la disponibilidad de puntos de suelo con una densidad variable, y la diversidad de cubiertas forestales en el área de estudio, permite analizar cómo estos factores afectan a la precisión de los MDE, utilizando alguno de los estadísticos del error, como por ejemplo el EMA (Bater and Coops 2009). Las categorías o grupos en los que se estratificó este error obtenido en los puntos de control seleccionados se describe en el Capítulo 3.

Es importante recalcar que estas técnicas estadísticas son aplicadas a un conjunto finito de puntos de control, por lo que no representan mediante una superficie continua cómo se distribuye el error que posee el MDE generado. Es necesaria, por tanto, la utilización de otro tipo de métodos que caractericen el error de forma espacial y los “árboles de decisión” pueden ser una buena alternativa para hacerlo. Un árbol de decisión es una representación gráfica con carácter analítico que informa y ayuda al usuario a tomar la decisión más acertada, desde un punto de vista probabilístico, ante un abanico de

posibilidades que tal vez no son evidentes con otros estadísticos más tradicionales. Es por ello que los árboles de decisión son considerados una técnica de “minería de datos” que prepara, sondea y explora los datos para extraer información “oculta” en ellos (Berlanga *et al.* 2013). Mediante su aplicación se pueden resolver problemas de predicción, clasificación y segmentación.

En términos generales, los árboles de decisión crean un modelo basado en diagramas de flujo jerárquico que muestran una serie de reglas sobre la decisión que se debe tomar en la asignación de un valor de salida a un determinado dato. Dicho de otra manera, clasifican casos en grupos o pronostican valores de una variable dependiente (criterio), basada en valores de variables independientes o predictoras. Básicamente, los árboles de decisión se componen de los siguientes elementos (IBM SPSS 2011, Berlanga *et al.* 2013) :

- Nodo de decisión: es aquel que indica que una decisión necesita tomarse en ese punto del proceso.
- Nodo de probabilidad: es el que indica que en ese punto del proceso ocurre un evento aleatorio.
- Nodo terminal: es aquel en el que todos los casos tienen el mismo valor para la variable dependiente, por lo que es considerado un nodo homogéneo o “puro” que no requiere ninguna división adicional.
- Rama: es el elemento que muestra las distintas vías que se pueden emprender cuando ocurre algún evento aleatorio o decisión y que tiene su final en un nodo puro.

La clasificación se realiza según una serie de preguntas sobre las variables predictoras, empezando por el nodo origen o raíz y continuando el camino determinado por las respuestas a las preguntas en los nodos internos, hasta llegar a un nodo terminal, conformándose así las reglas de decisión. La etiqueta asignada al nodo es la que determina la clasificación asignada. Para evitar el sobreajuste del modelo, se puede realizar una “poda” del árbol que elimine las ramas poco significativas, lo que puede dotar al árbol de mayor capacidad de generalización a costa de reducir el grado de pureza de sus nodos (Larose 2005). Hay que tener en cuenta que, en determinados casos, seguir haciendo particiones sucesivas hasta conseguir una clasificación pura, podría dar lugar a un árbol muy grande, con tantos nodos como datos hay en la muestra.

Existen diversos métodos de división para los árboles, entre los que destacan los siguientes, cuyas características se sintetizan a continuación de acuerdo con IBM SPSS (2011) y Berlanga *et al.* (2013):

- CHAID (*Chi-square automatic interaction detector*): consiste en un algoritmo que permite la detección automática de interacciones mediante Chi-cuadrado. En cada paso, CHAID elige la variable predictora que presenta la interacción más fuerte con la variable dependiente. Las categorías de cada predictor se funden si no son significativamente distintas respecto a la variable dependiente.
- CHAID exhaustivo: supone una modificación de CHAID que examina todas las divisiones posibles para cada predictor y trata todas las variables por igual, independientemente del tipo y el número de categorías.
- Árboles de clasificación y regresión (*CART-Classification and regression trees*) desarrollado por (Breiman *et al.* 1984): es un algoritmo de árbol binario completo (dos ramas por división) que genera subconjuntos precisos y homogéneos. CART divide los datos en segmentos para que sean lo más homogéneos posible respecto a la variable dependiente.
- QUEST (*Quick, unbiased, efficient, statistical tree*): es un algoritmo estadístico que crea árboles binarios precisos de forma rápida, evitando el sesgo que presentan otros métodos, al favorecer

los predictores con muchas categorías. Su limitación radica en que sólo puede utilizarse QUEST si la variable dependiente es nominal.

Las diferencias principales entre los distintos algoritmos de construcción de árboles de decisión radican en las estrategias de poda y en la regla adoptada para partir los nodos.

En el caso concreto del análisis del error atributivo de los MDE generados con los datos LiDAR-PNOA, se ha utilizado el análisis CART de acuerdo con la experiencia realizada por Bader and Coops (2009), así como por las ventajas que presenta su uso frente a otros métodos (Breiman *et al.* 1984, IBM SPSS 2011):

- Es un método no paramétrico.
- No requiere que las variables sean seleccionadas de antemano porque el algoritmo es capaz de identificar las más significativas y eliminar del análisis aquellas que no lo son (variables aleatorias).
- La estructura del árbol, es decir, el resultado, no se altera si una o varias variables independientes son transformadas a logaritmo o raíz cuadrada.
- Tiene una alta tolerancia a los valores atípicos (ruido de la muestra).
- Su carácter descriptivo posibilita la interpretación sencilla de las decisiones tomadas por el modelo, ya que se tiene acceso a las reglas de predicción, al contrario de lo que ocurre, por ejemplo, con las Redes Neuronales Artificiales. Este hecho, permite su implementación, mediante algebra de mapas, en un entorno SIG, para espacializar el resultado.
- Su conceptualización se aproxima al carácter de la geografía, ya que sintetiza reglas de decisión que generalizan un determinado fenómeno, en este caso, la incertidumbre en el error medio absoluto del MDE interpolado.

La variable que se quiere explicar, es decir, la variable dependiente continua seleccionada, fue el EMA, obtenido en cada uno de los puntos de control, mientras que las variables independientes eran aquellas que podían explicar mejor la variabilidad espacial del error: pendiente topográfica, densidad de puntos y tipo de cubierta de suelo. La construcción del árbol comienza en el nodo raíz, que incluye todos los puntos de la muestra, a partir del cual el algoritmo identifica la variable para dividirlo en 2 nodos hijos. El criterio de selección de las variables que van apareciendo sucesivamente en el proceso de creación del árbol se basa en una medida de pureza; la función de partición de Gini, que es la que más se utiliza para asegurar una pureza máxima en los nodos (IBM SPSS 2011).

2.2.5 Inventario forestal y estimación de variables dasométricas

Los datos de inventario forestal en las masas de *P. halepensis* necesarios para llevar a cabo una aproximación basada en áreas, en dos fases o de doble muestreo (Naesset 1997, 2002) fueron obtenidos en 45 parcelas circulares de 15 m de radio, distribuidas mediante un muestreo aleatorio estratificado, tratando de lograr un conjunto de datos representativo de la variabilidad de alturas de la masa forestal y de las pendientes topográficas existentes (Naesset and Okland 2002), dado que la realización de un muestreo sistemático incrementaría excesivamente el tamaño de la muestra, siendo prácticamente inabordable con los medios disponibles. Conviene recordar que el objetivo de este inventario de campo es cuantificar de forma precisa, sobre una pequeña fracción de la superficie total del área de estudio, ciertos parámetros dasométricos que constituyen las variables dependientes. En base a la información obtenida en las parcelas, se ajustarán los modelos de estimación para toda la superficie, por lo que el inventario debe recoger toda la variabilidad de ésta para minimizar los errores de extrapolación.

Una de las ventajas que presenta el inventario forestal con LiDAR es la sencilla estratificación de la masa. Para lograr unos resultados adecuados con un inventario tradicional, la superficie debe dividirse en estratos lo más homogéneos posibles en cuanto a especie y estructura de la masa, lo que implica un profundo conocimiento previo de la zona a inventariar. Sin embargo, los datos LiDAR proporcionan información sobre las variaciones en la estructura de la masa, y las variables o métricas derivadas de los retornos ya introducen dicha información en el modelo de estimación, por lo que sólo es necesario estratificar en función de la especie, que lógicamente presentará una distribución diferente de los retornos en función de su estructura. Por todo ello, en este trabajo se consideró únicamente el estrato arbóreo compuesto por masas monoespecíficas de *P. halepensis*, independientemente de su edad.

El número de las parcelas de inventario de campo se determinó en función de la validez estadística (Hair *et al.* 1999, Naesset 2002), siendo de forma circular para replantar tan solo una localización, correspondiente al centroide de la parcela, para lo cual se utilizó el equipo GNSS de Leica Geosystems® VIVA GS15-CS10. Respecto de las dimensiones, se optó por un diámetro de 30 m porque garantizaba un número mínimo de 250 retornos láser, necesarios para calcular las variables independientes en una fase posterior.

Dentro de cada parcela se midió la altura total y la altura de la copa viva de cada árbol utilizando un hipsómetro Vertex IV de Haglöf Sweden®, y el diámetro del tronco a la altura del pecho (dap), siempre que éste fuera mayor de 7,5 cm, con una forcípula Mantax Precision Blue de Haglöf Sweden® con graduación milimétrica (Figura 19). En cada localización se tomaron datos de diferentes factores relacionados con la vegetación y el estado fitosanitario de la masa, así como fotografías.



Figura 19. A la izquierda, Vertex IV, bastón extensible, transpondedor T3 y adaptador 360° de Haglöf Sweden®. El Vertex utiliza ultrasonidos para medir las distancias y asume automáticamente que el objeto de la medición se posiciona perpendicularmente al terreno. La altura se calcula con trigonometría al considerar el ángulo y la distancia. A la derecha, forcípula Mantax Precision Blue de Haglöf Sweden®.

Estos datos fueron el *input* necesario para calcular en cada parcela, la altura media (Hm, ecuación 7), el diámetro cuadrático medio (Dg, ecuación 8), el área basimétrica (G, ecuación 9), el volumen (V, ecuación 10) empleando la tarifa de cubicación del Tercer Inventario Forestal Nacional (IFN3) en la provincia de Zaragoza⁸, la densidad de pies (S, ecuación 11), la ratio de copa (CR, ecuación 12) y las diferentes fracciones de biomasa utilizando las ecuaciones alométricas para el *P. halepensis* desarrolladas por Ruiz-Peinado *et al.* (2011) (ecuaciones 13 a 19):

$$Hm (m) = \frac{\sum_{i=1}^N h_i}{N} \quad (7)$$

$$Dg (cm) = \sqrt{\frac{4 \cdot G}{S \cdot \pi}} \quad (8)$$

$$G (m^2) = \frac{\frac{\pi}{4} \sum dap_i^2}{a} \quad (9)$$

$$V (m^3) = \frac{\sum_{i=1}^N 0,00094 \cdot dap_i^{1,9209} \cdot h_i^{0,7264}}{a} \quad (10)$$

$$S (n^{\circ} de pies) = \frac{N}{a} \quad (11)$$

$$CR (\%) = \frac{L}{Hm} \cdot 100 \quad (12)$$

$$P_t = 0,0139 \cdot dap_i^2 \cdot h_i \quad (13)$$

$$P_{r7} (kg) = [3,926 \cdot (dap_i - 27,5)] \cdot Z; \text{ Si } dap_i \leq 27,5 \text{ cm entonces } Z = 0; \text{ Si } dap_i > 27,5 \text{ cm entonces } Z = 1 \quad (14)$$

$$P_{r2-7} (kg) = 4,257 + 0,00506 \cdot dap_i^2 \cdot h_i - 0,0722 \cdot dap_i \cdot h_i \quad (15)$$

$$P_{r2+a} (kg) = 6,197 + 0,00932 \cdot dap_i^2 \cdot h_i - 0,0686 \cdot dap_i \cdot h_i \quad (16)$$

$$P_{ra} (kg) = 0,0785 \cdot dap_i^2 \quad (17)$$

$$BA (kg) = P_t + P_{r7} + P_{r2-7} + P_{r2+a} \quad (18)$$

$$BT (kg) = BA + P_{ra} \quad (19)$$

Donde dap_i es el diámetro a la altura del pecho del árbol i medido en milímetros para el cálculo del V, en metros para el cálculo de G y en centímetros para el resto de variables dasométricas, a es la superficie de la parcela en hectáreas (0,070686 ha), h_i es la altura del árbol expresada en metros, N es el número total de pies de pino de la parcela, L es la longitud media de la copa viva en metros, P_t es el

⁸ Para la estimación de esta variable, se ha empleado la tarifa de cubicación del IFN3 en la provincia de Zaragoza para el *P. halepensis*. Pese a haber inventariado árboles tipo con parámetro de forma "1" (árboles fusiformes prácticamente en todo su fuste, con troncos maderables, limpios y derechos de más de 6 m, flecha inferior al 1% de su longitud, veta no torcida y $dap > 20$ cm), no se dispone para dicha especie y provincia de la tarifa de cubicación para la forma "1", por lo que se ha aplicado la tarifa para los árboles de forma "2" (árboles que cumplen las condiciones de ser fusiformes, tener troncos maderables de ≥ 4 m, ramificarse por la parte superior y no pertenecer a la forma 1).

peso de la biomasa de la fracción del tronco, P_{7-} es el peso de la biomasa de la fracción de ramas gruesas (diámetros >7 cm), P_{2-7} es el peso de la biomasa de la fracción de ramas medianas (diámetro entre 2 y 7 cm), $P_{<2}$ es el peso de la fracción de biomasa de ramas finas (diámetro <2 cm) y acículas, y P_{ra} es el peso de la biomasa de la fracción de raíces. De acuerdo con las ecuaciones 18 y 19, es posible calcular los kilogramos de BA y BT de cada pie de pino presente en una parcela, que son sumados para, con posterioridad, convertir la BA y la BT a kilogramos por hectárea.

La elección de estas variables estructurales responde a varios criterios (Bottalico *et al.* 2017): por su amplia utilización con fines de ordenación y gestión forestal, por su interés para obtener índices de complejidad estructural e indicadores de biodiversidad y por su uso en los inventarios forestales nacionales.

2.2.6 Estimación de la severidad post-incendio en campo

La medición de la severidad en campo evalúa el grado de pérdida o descomposición de la materia orgánica contenida en la vegetación y el suelo (Montorio *et al.* 2014). En este trabajo se ha utilizado el *Composite Burn Index* (CBI), un índice creado por Key and Benson (2003), que permite estimar el nivel de variación de las condiciones biofísicas de un sitio tras el incendio, considerando un conjunto de variables parametrizadas para evaluar el efecto del fuego en cinco estratos verticales organizados jerárquicamente para obtener valores promedio a nivel de sotobosque, dosel arbóreo y, sobre todo, a nivel global del sitio (Key and Benson 2006). Este enfoque multi-estrato proporciona una valoración más completa del efecto del fuego frente a otros métodos que se centran en cuestiones específicas, como por ejemplo, la proporción de copas consumida (Vega *et al.* 2008), el diámetro de las ramas supervivientes (Keeley *et al.* 2008) o la altura alcanzada por las llamas (Edwards *et al.* 2013).

El CBI es adecuado para su uso en una amplia variedad de escenarios post-fuego gracias a su diseño y operatividad, de ahí que sea uno de los índices más utilizados. Esto posibilita la comparación de comunidades vegetales que presentan diversas situaciones topográficas, climáticas y edáficas en distintos tipos de ecosistemas, composición de especies o estructuras de la vegetación (Amato *et al.* 2013). El índice incorpora estimaciones sobre factores tales como la condición y color del suelo, la cantidad de vegetación o de combustible consumido, el rebrote de las plantas quemadas, la aparición de nuevas especies, la altura de la quema, etc.

El CBI está diseñado para evaluar la severidad del fuego en la primera temporada de crecimiento después de un incendio, por lo que se visitaron los cuatro incendios de la zona de estudio C en los meses inmediatamente posteriores a su ocurrencia. Inicialmente, Tanase *et al.* (2011) establecieron 247 parcelas de muestreo repartidas entre los incendios de Zuera, Jaulín, Aliaga y Los Olmos, considerando desde zonas no quemadas hasta zonas de alta severidad. Sin embargo, las tareas de restauración y gestión forestal llevadas a cabo eliminaron la estructura vegetal existente en algunas de ellas, por lo que la muestra se redujo a 169 parcelas útiles para su asociación estadística con los datos LiDAR. En los dos incendios situados en el sector central del valle de Ebro (Zuera y Jaulín), área principal de estudio de esta tesis doctoral, se ubicaron casi el 60% de las parcelas. Cabe destacar que la mayor parte de los bosques fueron afectados por severidad alta, por lo cual, los puntos de muestreo no se distribuyen proporcionalmente por rangos de severidad.

De acuerdo con Tanase *et al.* (2011) y las indicaciones de Key and Benson (2006), las parcelas circulares de muestreo de 30 m de diámetro se establecieron en sitios con vegetación homogénea y pendiente constante; es decir, cada sitio de muestreo representaba zonas de condiciones semejantes respecto del tipo de vegetación y del nivel de severidad. En cada parcela se tomaron fotografías y el centro se georreferenció mediante un GPS centimétrico.

A la hora de llevar a cabo la estimación del estado de la vegetación tras el incendio, se examinó visualmente por estratos y de modo independiente, el grado de cambio con respecto al estado previo en una escala decimal de 0,0 (sin alteración) a 3,0 (totalmente alterada) con incrementos de 0,1. Concretamente, los cinco estratos tipo que consideran Key and Benson (2006) son:

1. Sustrato superficial del suelo.
2. Hierbas, arbustos y árboles de menos de 1 m de altura.
3. Arbustos y árboles de 1 a 5 m de altura.
4. Árboles intermedios de 10-25 cm de diámetro de tronco y 8-20 m de altura.
5. Grandes árboles (maduros, dominantes y codominantes del dosel superior) mayores que los anteriores y que reciben luz solar directa.

Finalmente, las puntuaciones de cada estrato fueron promediadas para obtener el valor de la severidad media (Key and Benson 2006). Los factores o los estratos que no se encontraron en la parcela se descartaron, así que las principales variables estimadas fueron: el consumo de combustible, la alteración del follaje, los cambios en la cubierta vegetal, la mortalidad de los árboles y la altura de la llama (Tanase *et al.* 2011).

2.2.7 Identificación en campo de combustibles forestales

En la literatura científica pueden encontrarse diferentes clasificaciones de tipos de combustibles basadas en las ecuaciones de propagación de Rothermel (1972). Una de las más conocidas es la propuesta por Albini (1976) para los Estados Unidos, denominada NFFL (*Northern Forest Fire Laboratory*), que se diseñó para simular incendios de superficie a través del programa informático BEHAVE. Así, en los modelos 1 a 3, el elemento propagador es el pasto, en los modelos 4 a 7 es el matorral, en los modelos 8 a 9 es la hojarasca bajo arbolado y en los modelos 10 a 13 son los restos de corta. A finales de la década de 1980, el ICONA (Instituto para la Conservación de la Naturaleza) realizó una traducción de los trece modelos que componen el NFFL para poderlos aplicar en España (Chuvieco and Martín 2004). Posteriormente, en el marco de un proyecto europeo (Prometheus 1999), se intentaron adaptar a las características particulares de la vegetación mediterránea, definiéndose siete tipologías (García *et al.* 2011, Tabla 6 y Figura 20). Este modelo *Prometheus* presenta un marcado carácter estructural a la hora de modelar el comportamiento del fuego, ya que tiene en cuenta la altura y la densidad de los elementos propagadores, que pueden agruparse en: pastos, matorral y hojarasca/restos de cortas (Arroyo *et al.* 2008, Ramírez *et al.* 2013). Por estas razones ha sido seleccionado para llevar a cabo la identificación de tipos de combustible en el campo. A pesar de que se trata de una tarea costosa en términos de tiempo y dinero, la principal ventaja del trabajo de campo es que el investigador está en contacto físico con el combustible, lo cual es necesario cuando existe una notable heterogeneidad de la vegetación y se requiere minimizar el error en la identificación para crear conjuntos de datos de referencia “verdad-terreno” que puedan ser relacionados con productos de teledetección (Arroyo *et al.* 2008) en una fase posterior de entrenamiento de una clasificación digital supervisada.

Entre julio y septiembre de 2014 se llevó a cabo una campaña de trabajo de campo con la finalidad de identificar *in situ* las siete tipologías de combustibles existentes según la clasificación del modelo *Prometheus* en el área de estudio B (Figura 20, Riaño *et al.* 2002):

- Tipo 1: pastizales agrícolas y vegetación herbácea con un porcentaje de cubrimiento >60%.
- Tipo 2: pastizales y arbustos pequeños de 0,3-0,6 m con un 30-40% de herbáceas. La cobertura de arbustos es >60% y la arbórea <50%.
- Tipo 3: arbustos medianos de 0,6-2,0 m, así como árboles jóvenes resultantes de una

- regeneración natural o una reforestación. La cobertura de arbustos es >60% y la arbórea <50%.
- Tipo 4: arbustos altos de 2,0-4,0 m y árboles regenerados. La cobertura de arbustos es >60% y la arbórea <50%.
 - Tipo 5: arbolado de más de 4 m sin sotobosque. La cobertura de arbustos <30% y la arbórea >50%.
 - Tipo 6: arbolado de más de 4 m con sotobosque, donde la distancia de éste con la base del dosel arbóreo es superior a 0,5 m. La cobertura de arbustos es >30% y la arbórea >50%.
 - Tipo 7: arbolado de más de 4 m con sotobosque alto y muy denso, donde la distancia de éste con la base del dosel arbóreo es inferior a 0,5 m. La cobertura de arbustos es >30% y la arbórea >50%.

Tabla 6. Descripción de los parámetros de carga de combustible, espesor del estrato de combustible y humedad de extinción de los tipos de combustible según *Prometheus*. La carga de combustible de materia muerta se divide en función del tiempo que tardan en alcanzar el equilibrio con la humedad ambiental: 1 h (diámetro de las partículas entre 0-0,62 cm), 10 h (entre 0,63-2,54 cm) y 100 h (entre 2,55-7,62 cm) (Chuvieco and Martín 2004).

Modelo	Carga de combustible (kg/m ²)				Espesor del estrato combustible (cm)	Humedad de extinción (%)
	1 h	10 h	100 h	Vivo		
1	0,166	0,000	0,000	0,000	30,48	30
2	0,225	0,112	0,000	0,449	60,96	30
3	0,360	0,070	0,000	1,350	107,16	30
4	1,123	0,898	0,449	1,123	182,88	30
5	0,740	0,100	0,020	0,000	5,69	30
6	0,449	0,225	0,112	0,112	30,48	30
7	0,540	0,110	0,030	0,720	100,55	30



Figura 20. Esquema utilizado para identificar en el campo los tipos de combustible según *Prometheus*. Las fotografías que se muestran corresponden a distintas localizaciones de la zona de estudio B.

Los datos fueron obtenidos en 108 parcelas circulares de 30 m de diámetro, distribuidas mediante un muestreo aleatorio estratificado, tratando de lograr una muestra representativa de la variabilidad de la cobertura vegetal (altura y fracción de cabida cubierta) y de las pendientes topográficas de la zona de estudio B. Las coordenadas del centro de cada parcela fueron tomadas con el equipo GNSS de Leica Geosystems® VIVA GS15-CS10. En cada una se hizo un inventario forestal (altura y diámetro de árboles, porcentaje de cobertura y altura de matorral) y se tomaron fotografías: cuatro desde el centro y en dirección a los puntos cardinales norte, sur, este y oeste, y 2-4 fotos desde el perímetro de la parcela hacia el interior.

2.2.8 Imágenes ópticas e índices espectrales

La tesis doctoral se fundamenta en el uso de datos LiDAR; sin embargo, se han empleado otras fuentes de información de teledetección óptica para enriquecer el análisis de la estimación de la severidad y la cartografía de tipos de combustible forestal. Tal como se puso de manifiesto en el Capítulo 1, la combinación de datos permite aprovechar las ventajas propias de cada tipo para lograr resultados satisfactorios.

Imágenes Landsat TM

La acción del fuego sobre la vegetación produce cambios en su composición, densidad, vigor y color, así como en el contenido de humedad de ésta. En este sentido, las regiones infrarrojas del espectro electromagnético, que son registradas por los sensores ópticos multiespectrales de teledetección, son muy útiles para conocer el impacto del fuego. Concretamente, el programa Landsat es la única fuente de información temporal y espacial coherente y continua a nivel global desde 1972, dentro de la categoría de sensores ópticos de media-alta resolución espacial. Esto permite comparar información previa y posterior al fuego para la evaluación de la magnitud del cambio causado en la vegetación.

Por tales razones, las imágenes ópticas utilizadas en esta investigación provienen del satélite Landsat TM 5, que opera en siete bandas espectrales diferentes, tiene una resolución espacial de 30 m en todas las bandas excepto en la 6 y una resolución temporal de 16 días. De estas siete bandas, son especialmente valiosas para el análisis de la severidad post-fuego la banda 4 o banda del infrarrojo cercano (NIR), que se registra en las longitudes de onda entre 0,76 μm y 0,90 μm , y la banda 7, una de las dos del infrarrojo medio de onda corta (SWIR), situada entre 2,08 μm y 2,35 μm . La banda 4 depende principalmente del índice de refracción de la morfología foliar y de las discontinuidades dentro de la hoja, mientras que la banda 7 es sensible al contenido de agua, tanto de la vegetación como del suelo (Soverel *et al.*, 2010).

El conjunto de imágenes ópticas analizadas (columna 199; fila 031 y 032) cubre los incendios englobados en el área de estudio C. En la Tabla 7 se proporciona la fecha de adquisición de cada imagen junto con la del incendio forestal correspondiente.

Tabla 7. Imágenes Landsat utilizadas.

Zona de estudio	Fecha de inicio del incendio	Imagen pre-fuego	Imagen post-fuego	Columna/Fila
Zuera	06/08/2008	21/07/2008	22/08/2008	199/31
Jaulín	29/07/2009	22/06/2009	26/09/2009	199/31
Aliaga	22/07/2009	22/06/2009	26/09/2009	199/31
Los Olmos	22/07/2009	22/06/2009	26/09/2009	199/32

La elección de estas imágenes cumple con la necesidad de obtener datos con un mínimo espaciamiento temporal, ya que la variabilidad de la respuesta espectral de la vegetación no quemada se reduce al mínimo. Dichas imágenes fueron proporcionadas por el PNT con un nivel de procesamiento mínimo que consiste en la eliminación de los errores sistemáticos de calibración y de localización de los píxeles en una matriz regular. Las imágenes estaban proyectadas en un sistema de proyección Universal Transverse Mercator (UTM), utilizándose la información sobre la posición de la plataforma en el momento de la adquisición. El error sistemático residual de localización se situó en torno a 250 m y las imágenes resultantes fueron proporcionadas sin ningún tipo de corrección atmosférica. Por consiguiente, la eliminación de los errores sistemáticos por parte del proveedor hizo innecesaria la aplicación de un pre-tratamiento de restauración de píxeles o líneas perdidas. Por el contrario, sí que fue ineludible realizar una corrección geométrica, seguida de una corrección radiométrica. La corrección geométrica de las imágenes se basó en la toma de 50 puntos de control sobre la referencia de las ortofotografías del PNOA de 0,5 m de resolución. Teniendo en cuenta la complejidad topográfica del área de estudio, para la modelización de la transformación se utilizó un polinomio de segundo orden. La bondad del ajuste presentó un RMSE <0,65 píxeles. Se utilizó la misma proyección cartográfica que tienen las parcelas de campo (UTM, zona 30 N, datum ETRS89), dado que iban a ser analizadas conjuntamente. La corrección radiométrica se llevó a cabo con un algoritmo implementado en la mayoría de los paquetes de programas comerciales que trabaja con un repertorio de condiciones atmosféricas (en función de la latitud, la altitud y los tipos de aerosoles) y de valores de visibilidad atmosférica (Tanase *et al.*, 2011).

Imagen SPOT 5

La caracterización espectral de los combustibles forestales localizados en el área de estudio B, se ha llevado a cabo empleando una imagen obtenida por el satélite SPOT 5 (columna 038/fila 266) de 10 m de resolución espacial, cuyas bandas son: verde (0,50-0,59 μm), rojo (0,61-0,68 μm), infrarrojo cercano (0,78-0,89 μm) e infrarrojo medio de onda corta (1,58-1,75 μm).

La escena, de 60 x 60 km, fue registrada el 28 de agosto de 2010, aproximadamente seis meses antes de la captura de las nubes de puntos LiDAR del PNOA. Se consideró aceptable el desfase temporal entre ambas fuentes de información, ya que el paisaje vegetal del área de estudio está dominado por especies perennifolias. Además, se hizo uso del repositorio de imágenes gratuitas y disponibles en el PNT.

La corrección geométrica de la imagen se basó en la toma de 33 puntos de control sobre la referencia de las ortofotografías del PNOA de 0,5 m de resolución, tanto en el interior como en el entorno del CENAD “San Gregorio”, garantizando así la correcta georreferenciación de esta área (RMSE <0,5 píxeles). Se utilizó la misma proyección cartográfica que tienen las nubes de puntos LiDAR del PNOA (UTM, zona 30 N, datum ETRS89), dado que iban a ser analizadas conjuntamente.

La corrección radiométrica y la normalización topográfica conllevaron el uso de un MDE de 10 m de resolución espacial, remuestreado del MDE generado por el PNOA-LiDAR (5 m de resolución). Para llevar a cabo esta tarea de normalización topográfica de la imagen, se empleó el método incluido en el *software* ERDAS IMAGINE 2013 (Hodgson and Shelley 1994), obteniendo los coeficientes de calibración de cada banda del fichero de cabecera de la propia imagen, así como el ángulo de elevación solar (55.093887°) y el ángulo acimutal solar (151.657861°).

Generación de índices espectrales

Las bandas espectrales registradas por los sensores ópticos presentan un comportamiento diferente según las características de la superficie terrestre (contenido de humedad, estructura de la vegetación,

actividad fotosintética y estado fenológico de la vegetación, composición mineral del suelo, etc.). Estudios previos de diversa índole, entre los que se encuentran los relacionados con la estimación de variables como la severidad del fuego o la clasificación digital de combustibles forestales, se han apoyado en el uso, tanto de los valores de reflectividad de las bandas originales, como de transformaciones de éstas para obtener nuevas bandas (neocanales) que ligen las bandas originales mediante algoritmos matemáticos con objeto de extraer información aún más relevante sobre las propiedades de las superficies de la escena (Chuvieco 2010).

En concreto, uno de los índices más habituales para la estimación de la severidad post-incendio mediante sensores ópticos, como Landsat TM y ETM+, se conoce como *Normalized Burn Ratio* (NBR), el cual relaciona la banda 4 del infrarrojo próximo (entre 0,7 y 1,3 μm) y la banda 7 del infrarrojo medio de onda corta (entre 1,3 y 2,5 μm). El infrarrojo próximo está directamente relacionado con la superficie foliar y la actividad fotosintética, mientras que el infrarrojo medio de onda corta responde positivamente a la escasez de humedad del suelo y de los combustibles quemados. Por tanto, el NBR es positivo en las zonas con actividad fotosintética intensa y negativo en zonas de baja productividad vegetal o sin vegetación (Key and Benson 2006, ecuación 20):

$$NBR = \frac{(R_4 - R_7)}{(R_4 + R_7)} \quad (20)$$

Donde R_4 y R_7 son las reflectividades de la bandas 4 y 7 del sensor Landsat TM.

Para discriminar las zonas quemadas de las no quemadas y proporcionar una medida cuantitativa del cambio ocurrido en el área, el NBR post-incendio se resta al NBR pre-incendio (ecuación 21), dando lugar al índice sintético *Differenced* o *Delta NBR* (dNBR):

$$dNBR = NBR_{pre-incendio} - NBR_{post-incendio} \quad (21)$$

El dNBR (multiplicado por 10^3) provee un rango continuo de valores válidos que oscilan entre -550 y 1.350. Los valores negativos suelen ser el resultado de la presencia de nubes en la imagen pre-fuego o de la regeneración vegetal rápida (herbáceas) en la imagen post-fuego ($-500 < dNBR < -100$). Los valores positivos (entre 100 y 1.300) se producen por el grado de impacto del fuego sobre la vegetación y el suelo, el cual puede discretizarse en diferentes clases de severidad: no quemado de -101 hasta +99; baja de +100 hasta +269; baja-moderada de +270 hasta +439; moderada-alta de +440 hasta +659; y alta de +660 hasta +1.300. Valores por encima de 1.350 se corresponden normalmente a zonas nubosas. No obstante, dichos umbrales comúnmente empleados, pueden variar en función del estado fenológico de la vegetación, del ángulo de iluminación solar y del lapso de tiempo entre el momento del incendio y el de estimación de la severidad (Key and Benson 2006, Chuvieco 2010).

Por otro lado, una versión alternativa del anterior índice espectral es el *Relative dNBR* (RdNBR), propuesto por Miller and Thode (2007), para convertir el cambio absoluto medido por el dNBR en una medida relativa que varíe linealmente con la severidad (ecuación 22).

$$RdNBR = \frac{NBR_{pre-incendio} - NBR_{post-incendio}}{\sqrt{|NBR_{pre-incendio}|}} \quad (22)$$

Miller and Thode (2007) señalaron dos ventajas principales del índice RdNBR frente al dNBR. La primera, es que proporciona una aproximación más consistente sobre el impacto del fuego en el ecosistema, posibilitando la comparación de valores en diferentes escenarios espaciales y temporales al eliminarse el sesgo de las condiciones de la vegetación pre-fuego (Miller *et al.* 2009). La segunda, es que

mejora la precisión en la identificación de las zonas con altas severidades, especialmente cuando la vegetación previa al incendio es heterogénea, como suele ser habitual en los espacios forestales mediterráneos. Por consiguiente, tal como se describe en el Capítulo 3 de la tesis, se ha asumido que los cambios que evidencian los valores de dNBR y RdNBR se correlacionan con los cambios ambientales provocados por el fuego y estimados en campo a través del índice CBI.

Por otro lado, en el contexto de la clasificación digital de tipos de combustibles forestales, se han generado otro tipo de neocanales relacionados con índices de vegetación, que en su mayoría, se fundamentan en el contraste espectral entre la longitud de onda del rojo (entre 0,6 y 0,7 μm) y la del infrarrojo cercano (entre 0,7 y 1,3 μm). Esto responde a que la vegetación sana o verde presenta un mínimo relativo de reflectividad en el rojo, debido a que los pigmentos de la hoja absorben la mayor parte de la energía que reciben, y un máximo absoluto en el infrarrojo cercano, como consecuencia de la dispersión de la radiación que ejerce el interior de la estructura celular. Cuanto mayor sea la diferencia entre la reflectividad de estas dos bandas, mayor vigor vegetal presentará la cubierta observada, mientras que bajos contrastes son sinónimos de una vegetación con poca densidad, senescente o enferma (Chuvieco 2010).

En este principio teórico se apoyan la mayor parte de los índices de vegetación existentes, entre los que destaca por su alta usabilidad para estimar diversos parámetros de la cubierta vegetal (contenido de clorofila, estrés hídrico, flujo neto de CO_2 , radiación solar absorbida fotosintéticamente, productividad, índice de área foliar, dinámica fenológica, etc.), el Índice de Vegetación de Diferencia Normalizada (NDVI-*Normalized Difference Vegetation Index*, ecuación 23) (Chuvieco 2010).

$$NDVI = \frac{\rho_{i \text{ infrarrojo cercano}} - \rho_{i \text{ rojo}}}{\rho_{i \text{ infrarrojo cercano}} + \rho_{i \text{ rojo}}} \quad (23)$$

Donde $\rho_{i \text{ infrarrojo cercano}}$ y $\rho_{i \text{ rojo}}$ indican las reflectividades del píxel i en la banda del infrarrojo cercano y del rojo, respectivamente. El rango de valores del NDVI varía entre -1 y +1, pudiéndose señalar como umbral crítico para cubiertas vegetales un valor de NDVI en torno a 0,1 y para la vegetación densa entre 0,5 y 0,7.

Por otro lado, para analizar el contenido de humedad de la vegetación, las bandas del infrarrojo medio de onda corta, comprendido entre 1,3 y 2,5 μm , y del infrarrojo cercano, son las más idóneas para este propósito, ya que al aumentar el contenido de agua en el sustrato o en la vegetación disminuye paralelamente la reflectividad en el SWIR, mientras que al mismo tiempo, el infrarrojo cercano informa sobre el grado de deterioro de la hoja.

Entre los diversos índices propuestos que contrastan las dos bandas citadas, se ha utilizado el NDII (*Normalized Difference Infrared Index*), definido por Hunt and Rock (1989) como (ecuación 24):

$$NDII = \frac{\rho_{i \text{ infrarrojo cercano}} - \rho_{i \text{ SWIR}}}{\rho_{i \text{ infrarrojo cercano}} + \rho_{i \text{ SWIR}}} \quad (24)$$

Donde $\rho_{i \text{ infrarrojo cercano}}$ y $\rho_{i \text{ SWIR}}$ son las reflectividades del infrarrojo cercano y del SWIR, respectivamente.

2.2.9 Variables derivadas de las nubes de puntos LiDAR

Para ajustar los modelos de regresión, se relacionan las variables LiDAR o independientes (también denominadas métricas), con las variables estimadas o medidas en el campo. Es por ello, que las variables independientes se calculan en el área de la nube de puntos que corresponde espacialmente a cada una de las parcelas de campo.

Nótese que el procedimiento para obtener las variables LiDAR comienza con el filtrado de la nube de puntos (utilizando el algoritmo que mejor resultado ha proporcionado), continúa con la generación de un MDE para la zona de estudio (utilizando el método que minimiza el error) y finaliza implementando los comandos “ClipData” y “CloudMetrics” incluidos en el *software* FUSION LDV 3.30 (McGaughey 2009). En este último paso, la nube de puntos es “recortada” con las dimensiones de las parcelas circulares de campo y las alturas de los retornos láser son normalizadas empleando el MDE, lo que permite calcular, para cada parcela, un conjunto de variables relacionadas con las alturas reales de los retornos láser y que comúnmente son utilizadas en el modelado de la estructura de la vegetación. En ocasiones, con el objeto de excluir de las métricas LiDAR los retornos procedentes del suelo y del sotobosque, conviene aplicar un umbral de altura sobre el terreno para la selección de los puntos utilizados en el cálculo (Naesset and Okland 2002). Así, para el inventario forestal que se ha llevado a cabo en la zona de estudio B, y que se describe en detalle en el Capítulo 3, se ha definido un umbral de 1 m para excluir los retornos que quedan por debajo de éste.

En la Tabla 8 se describen las variables LiDAR extraídas, susceptibles ser utilizadas en los modelos predictores, las cuales presentan una relación intuitiva con la estructura de la vegetación (Evans *et al.* 2009, McGaughey 2009).

Tabla 8. Variables derivadas de la distribución de las alturas de los retornos láser, donde x_i es el valor de la observación (altura del retorno), N es el número total de observaciones, r_i es el retorno, y pli es el pulso.

Variable	Descripción
Percentiles 1, 5, 10, 20, 25, 30, 40, 50, 60, 70, 75, 80, 90, 95 y 99 (P_{01} , P_{05} , P_{10} , etc.)	<p>Los diferentes valores de percentiles fueron calculados utilizando la siguiente metodología:</p> $(N - 1)P = I + d \begin{cases} I \text{ es la parte entera de } (N - 1)P \\ d \text{ es la parte decimal de } (N - 1)P \end{cases}$ <p>Donde N es el número de observaciones y P es el percentil dividido por 100.</p> <p>Si $d = 0$ entonces $\text{Percentil} = x_{i+1}$</p> <p>Si $d > 0$ entonces $\text{Percentil} = x_{i+1} + d(x_{i+2} - x_{i+1})$</p> <p>Donde x_i es el valor de la observación teniendo en cuenta que las observaciones están ordenadas de manera ascendente.</p>
Altura mínima	$x_i \text{ mínima}$
Altura máxima	$x_i \text{ máxima}$
Rango	$x_i \text{ máxima} - x_i \text{ mínima}$
Altura media (μ)	$\frac{\sum_{i=1}^N x_i}{N}$
Moda de las alturas	Valor x_i más frecuente en la parcela
Ratio de apertura del dosel (<i>canopy relief ratio</i>)	$\frac{\mu - x_i \text{ mínima}}{x_i \text{ máxima} - x_i \text{ mínima}}$

Tabla 8. Continuación.

Variable	Descripción
Desviación típica (σ)	$\sqrt{\frac{\sum_{i=1}^N (x_i - \mu)^2}{N}}$
Varianza (σ^2)	$\frac{\sum_{i=1}^N (x_i - \mu)^2}{N}$
Coefficiente de variación	$\frac{\sigma}{\mu} 100$
Desviación absoluta media	$\frac{\sum_{i=1}^N (x_i - \mu)}{N}$
Rango intercuartil	$[P_{75}(x) - P_{25}(x)]$
Asimetría	$\frac{\sum_{i=1}^N (x_i - \mu)^3}{(N - 1)\sigma^3}$
Curtosis	$\frac{\sum_{i=1}^N (x_i - \mu)^4}{(N - 1)\sigma^4}$
Porcentaje de retornos que son primeros, segundos, terceros y cuartos	$\frac{\sum_{i=1}^N r_{i \text{ primeros}}}{N} 100; \frac{\sum_{i=1}^N r_{i \text{ segundos}}}{N} 100; \text{etc.}$
Porcentaje de retornos de suelo	$\frac{\sum_{i=1}^N r_{i \text{ suelo}}}{N} 100$
Porcentaje de retornos de vegetación	$\frac{\sum_{i=1}^N r_{i \text{ vegetación}}}{N} 100$
Porcentaje de pulsos con uno, dos o tres retornos	$\frac{\sum_{i=1}^N pl_{i \text{ con un retorno}}}{N} 100; \frac{\sum_{i=1}^N pl_{i \text{ con dos retorno}}}{N} 100, \text{etc.}$
Porcentaje de primeros retornos por encima de un umbral de altura determinado, de la altura media o de la moda (estimación de la fracción de cabida cubierta)	$\frac{\sum_{i=1}^N r_{i \text{ primeros}} > \text{umbral}}{\sum_{i=1}^N r_{i \text{ primeros}}} 100$
Porcentaje de retornos total por encima de un umbral de altura determinado, de la altura media o de la moda	$\frac{\sum_{i=1}^N r_i > \text{umbral}}{N} 100$
Porcentaje de retornos total por encima de un umbral de altura determinado, de la altura media o de la moda, sobre el total de primeros retornos	$\frac{\sum_{i=1}^N r_i > \text{umbral}}{\sum_{i=1}^N r_{i \text{ primeros}}} 100$
Porcentaje de retornos total dentro de un rango determinado de altura (densidad estratificada)	$\frac{\sum_{i=1}^N r_i > x_{i1} \text{ y } < x_{i2}}{N} 100$

2.2.10 Modelos de regresión

Modelo de regresión lineal múltiple

Para determinar si existe relación o independencia entre las variables no basta con el diagrama de dispersión, sino que es preciso llevar a cabo algún tipo de prueba de significación o índice (coeficientes de correlación) que permitan su cuantificación. El objetivo es construir modelos que expliquen el comportamiento de una variable de interés, como resultado del efecto de un conjunto de variables explicativas, por lo que es necesario, ante la gran cantidad de variables independientes candidatas, hacer un análisis previo a modo de preselección.

Se entiende que dos variables están relacionadas cuando ambas varían de forma conjunta. Para determinar la magnitud se dispone de una serie de índices que se emplean según sea el tipo de variables sobre las que se cuantifica la relación, como por ejemplo, el coeficiente de correlación de Pearson, el de Tau-b de Kendall o el de Spearman. Este último es el que más se ha venido utilizando en aplicaciones forestales, dado que puede utilizarse como una alternativa al de Pearson cuando las variables estudiadas incumplen el supuesto de normalidad. El coeficiente de correlación *rho* de Spearman es similar al coeficiente de correlación de Pearson, tomando valores entre -1 y 1, siendo el valor de 1 indicador de una relación lineal perfecta positiva o directa y el valor de -1 sinónimo de una relación lineal perfecta negativa o inversa (en ambos casos los puntos del correspondiente diagrama de dispersión se encuentran dispuestos en una línea recta). Por el contrario, el valor 0 indica relación lineal nula (Pardo and Ruíz 2005).

Si el valor y la significación del coeficiente de correlación permiten determinar que dos variables están correlacionadas linealmente, es razonable pensar que si se conoce el comportamiento de una variable se pueda predecir el comportamiento de la otra. La técnica estadística que permite explorar y cuantificar la relación entre una variable dependiente y una o más variables predictoras o independientes, es el análisis de regresión lineal simple o múltiple, respectivamente.

En el caso de un análisis de regresión lineal múltiple, lo que se construye como mejor predicción no es una recta sino un hiperplano. La estructura de la ecuación del modelo es la misma que la de la regresión lineal simple, sólo que con más predictores, cuyos coeficientes de correlación se calculan de modo que hagan mínimas las diferencias cuadráticas entre los valores observados y los pronosticados por el modelo (ecuación 25):

$$Y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \dots + \beta_n X_n \quad \text{expresado en puntuaciones directas} \quad (25)$$

Donde Y es la variable dependiente que se quiere predecir, β_0 es el valor de la constante, β_n es el coeficiente de regresión de la variable independiente X_n .

Cuando se dispone de varias variables independientes, existen dos formas de especificar el modelo de regresión: mediante especificación confirmatoria o mediante métodos de búsqueda secuencial. El primero es útil cuando se quiere emplear el modelo de regresión desde una perspectiva confirmatoria, siendo el investigador quien especifica las variables independientes a incluir (teniendo en cuenta que el conjunto de variables escogidas consigue la máxima predicción mientras se mantiene la parsimonia del modelo). Por su parte, los métodos de búsqueda secuencial sirven para elegir entre el conjunto de variables independientes, aquéllas que explican mejor la dependiente, por lo que tiene un carácter más exploratorio. Dado que en este trabajo se pretende seleccionar, del conjunto de variables independientes, las que mejor estimen la variable dependiente de interés, se ha realizado un análisis de correlación previo para hacer una primera selección de variables independientes a introducir en el modelo, y a continuación se ha hecho uso de un método de búsqueda secuencial por ser la opción más adecuada. Son tres los métodos de búsqueda secuencial más habituales: hacia delante (*forward*), hacia

atrás (*backward*) y de pasos sucesivos (*stepwise*). De éstos, el de pasos sucesivos es el más completo para seleccionar variables independientes en los métodos multivariantes (Hair *et al.* 1999).

Tras el análisis de regresión, se verifica su validez estadística garantizando que se cumplen con los principios que confirman las hipótesis básicas propias de un modelo de estas características (Hair *et al.* 1999, Pardo and Ruíz 2005):

- Principio de normalidad de los residuos tipificados, los cuales se distribuyen normalmente con media cero. La normalidad se puede contrastar mediante métodos gráficos y test estadísticos. Dentro de los primeros se encuentran el análisis del histograma y del gráfico de normalidad; destacando dentro de los segundos los test de simetría, la curtosis y las pruebas de Kolmogorov-Smirnov y Shapiro-Wilk.
- Principio de homocedasticidad, que se refiere a que la varianza de los residuos debe ser constante para cada valor de la variable independiente. Se puede evaluar este supuesto mediante el test de Breusch-Pagan. El gráfico de dispersión que relaciona los residuos tipificados frente a los valores pronosticados permite también realizar un diagnóstico visual de la homocedasticidad.
- Principio de linealidad entre la variable dependiente y las independientes, de carácter fundamental, ya que, si no se cumple, no es apropiado utilizar el método de regresión lineal. La forma más usual de evaluar la linealidad entre variables es el gráfico de dispersión, mediante el uso de los valores tipificados de los residuos y de los valores pronosticados. A nivel cuantitativo, el test RESET de Ramsey, sirve para evaluar la relación lineal entre las variables y detectar errores de especificación en sentido amplio.
- Principio de independencia, referido a que los residuos son independientes entre sí, es decir, los residuos constituyen una variable aleatoria. Para identificar este hecho se suele utilizar el gráfico de dispersión de los residuos tipificados frente a los valores pronosticados. Otra forma de evaluación es mediante la utilización del estadístico Durbin-Watson.
- Principio de no-colinealidad, que se refiere a la inexistencia de altas correlaciones entre las variables independientes utilizadas en el modelo, es decir, no existe relación lineal exacta entre ninguna de las variables independientes. El método más simple de identificar la colinealidad es el examen de la matriz de correlación de las variables independientes. Para evaluar la colinealidad del modelo de regresión ya ajustado se pueden utilizar estadísticos como el factor de inflación de la varianza (FIV).
- Principio de ausencia de casos influyentes o valores atípicos, que afectan notablemente al valor de la ecuación del modelo sesgándola, por lo que es conveniente verificar este supuesto a través de estadísticos, como el test de Bonferroni o el análisis de las distancias que expresan el grado en que cada caso se aleja del resto, por ejemplo a través de las distancias de Cook.

Aparte de estos seis supuestos, hay que tener en cuenta la influencia del tamaño de la muestra en la generalización de los resultados. Así, con el fin de que el modelo no estuviera sobre-ajustado a la muestra y fuera extrapolable, se siguió a Hair *et al.* (1999), que recomiendan entre 15 y 20 observaciones por cada variable independiente incluida en el modelo de regresión.

Por otro lado, además de utilizar las variables originales, y sobre todo si se incumple alguno de los principios anteriores –especialmente el de normalidad, homocedasticidad y linealidad–, la transformación logarítmica de las variables, tanto de la dependiente como de las predictoras, puede ser útil para resolver dichos problemas con las hipótesis básicas y para mejorar la significancia y los estadísticos que indican la bondad del ajuste (Means *et al.* 2000, Naesset 2002).

Cuando la relación entre las variables es exponencial, se puede ajustar una regresión lineal mediante la aplicación del logaritmo natural a la variable dependiente. El modelo lineal resultante, al cual se le aplica la transformación logarítmica inversa (conversión exponencial), es equivalente a un modelo exponencial en el que intervienen las variables originales (ecuación 26) (Pardo and Ruíz 2005).

$$Y = \beta_0 e^{\beta_1 X_1} e^{\beta_2 X_2} \dots e^{\beta_n X_n} FC \quad (26)$$

Cuando la relación entre las variables es potencial, se puede ajustar una regresión lineal mediante la aplicación del logaritmo natural tanto a las variables dependientes como a las variables independientes. El modelo lineal resultante, al cual se le aplica la transformación logarítmica inversa, es equivalente a un modelo potencial en el que intervienen las variables originales (ecuación 27) (Pardo and Ruíz 2005).

$$Y = \beta_0 X_1^{\beta_1} X_2^{\beta_2} \dots X_n^{\beta_n} FC \quad (27)$$

Como se puede observar en ambas ecuaciones, FC hace referencia al factor de corrección por el cual deben multiplicarse los modelos que han sufrido una transformación logarítmica inversa (ecuación 28), dado que se produce un sesgo en el modelo como consecuencia de dicha transformación (Baskerville 1972). El FC se calcula a partir del error estándar de estimación (SEE) (ecuación 29) de acuerdo con Sprugel (1983).

$$FC = e^{\frac{SEE^2}{2}} \quad (28)$$

$$SEE = \sqrt{\frac{\sum_{i=1}^N (Y_i - Y'_i)^2}{N - p}} \quad (29)$$

Donde Y_i es el valor real para la observación i , Y'_i es el valor pronosticado para la observación i , N es el número de observaciones y p es el número de parámetros del modelo.

Para cuantificar la proporción de varianza de la variable dependiente que es explicada por las independientes se utiliza el coeficiente de determinación, que es el cociente entre la suma de los cuadrados de la regresión y la suma total de cuadrados (ecuación 30).

$$R^2 = \frac{\sum_{i=1}^N (Y'_i - \bar{Y})^2}{\sum_{i=1}^N (Y_i - \bar{Y})^2} 100 \quad (30)$$

Donde R^2 es el coeficiente de determinación, Y_i es el valor real para la observación i , \bar{Y} es la media de los valores reales, e Y'_i es el valor pronosticado para la observación i . El valor de R^2 varía entre 0 y 1, siendo mayor el poder explicativo del modelo de regresión cuanto más alto sea (Hair *et al.* 1999).

La diferencia entre los valores reales de la variable dependiente Y_i y los predichos Y'_i mediante la ecuación lineal ajustada, recibe el nombre de residuo r_i y puede utilizarse para valorar globalmente la exactitud de la predicción. Asimismo, se ha seleccionado el RMSE por ser muy utilizado en los trabajos de estimación de variables dasométricas y porque su interpretación es sencilla: cuanto menor es el valor obtenido de RMSE, mejor es el modelo de regresión (ecuación 31).

$$RMSE = \sqrt{\frac{\sum_{i=1}^N (Y_i - Y'_i)^2}{N}} \quad (31)$$

Donde Y_i es el valor real para la observación i , Y'_i es el valor pronosticado para la observación i y N es el número de casos u observaciones.

Por último, se ha considerado el sesgo reportado por el modelo, definido como la media aritmética de los residuos (ecuación 32).

$$Sesgo = \frac{\sum_{i=1}^N (Y_i - Y'_i)}{N} \quad (32)$$

Una vez seleccionado el modelo de regresión más adecuado, hay que realizar un proceso de validación. Si la muestra es reducida, la mejor opción es utilizar la técnica de validación cruzada dejando un dato fuera (LOOCV - *Leave one-out cross-validation*). Se basa en eliminar un caso de la muestra –una parcela– en cada interacción y estimar su valor mediante el modelo especificado utilizando el resto de los casos de la muestra. De esta forma se realizan tantas iteraciones como casos tenga el conjunto de datos, a partir de las cuales se pueden calcular los valores de R^2 , RMSE y sesgo cada vez, así como los coeficientes obtenidos en cada iteración. Realizando un promedio de estos valores se obtiene el R^2 , el RMSE y el sesgo tras la validación, que se comparan con los del modelo inicial. Si los resultados son similares entre los estadísticos de bondad del ajuste pre y post validación, el modelo es aceptado y puede ser aplicado a toda el área de estudio (Bouvier *et al.* 2015).

Modelo de regresión logística

La regresión logística se utiliza para estimar la probabilidad de ocurrencia de una variable dependiente categórica binaria (dos grupos, 0/1), permitiendo la entrada de variables explicativas continuas y/o categóricas que no tienen por qué cumplir el supuesto de normalidad. Como se puede apreciar en la ecuación 33, se trata de una relación multivariante con coeficientes que indican la influencia relativa de las variables independientes (Hair *et al.* 1999, Menard 2010):

$$P(Y = 1) = \frac{\exp(\beta_0 + \beta_1 X_1 + \beta_2 X_2 + \dots + \beta_n X_n)}{1 + \exp(\beta_0 + \beta_1 X_1 + \beta_2 X_2 + \dots + \beta_n X_n)} \quad (33)$$

Donde $P(Y = 1)$ es la probabilidad de que la variable dependiente Y sea igual a 1, β_0 es la constante y β_n es el coeficiente de la variable independiente X_n .

Su naturaleza no lineal requiere un método de estimación diferente al de mínimos cuadrados, por lo que se utiliza el criterio de la máxima verosimilitud, que busca la estimación más probable de los coeficientes (Menard 2010). El resultado es una función continua cuyos valores de salida oscilan entre 0 y 1, de tal manera que los valores $<0,5$ se corresponden con la no ocurrencia de la variable dependiente y los valores $>0,5$ con la ocurrencia. Los coeficientes estimados (β_n) son medidas de los cambios en el ratio de probabilidades (*odds ratio*), de tal manera que un coeficiente positivo aumenta la probabilidad de ocurrencia, mientras que uno negativo la disminuye (Hair *et al.* 1999).

Como en la regresión lineal, el método aplicado para el ajuste de los modelos logísticos ha sido el de búsqueda secuencial por pasos sucesivos. No obstante, también se ha llevado a cabo un análisis exploratorio desde el enfoque de la especificación confirmatoria, empleando técnicas no paramétricas, como el coeficiente de correlación de Spearman y el test de Kruskal-Wallis, con objeto de conocer qué variables LiDAR son las más susceptibles de ser incluidas en el proceso de ajuste del modelo de regresión logística para estimar la severidad de los incendios de la zona de estudio C.

La medida global de cómo se ajusta el modelo viene dada por el valor de la verosimilitud (-2LL), que queda definido por la siguiente expresión (ecuación 34):

$$-2LL = -2\ln V \quad (34)$$

Donde $-2LL$ es el valor de la verosimilitud y V es el producto de todas las probabilidades predichas por el modelo de que los casos (parcelas de muestreo de campo) de la muestra tengan la condición (severidad) que realmente tienen. El valor mínimo para $-2LL$ es 0, de manera que cuanto más pequeño sea el valor de este parámetro, mejor será el ajuste (Hair *et al.* 1999).

En esta misma línea, el test de Hosmer-Lemeshow proporciona una medida final del ajuste que se basa también en la predicción real de la variable dependiente. Para ello, primeramente se dividen los casos en 10 clases aproximadamente iguales, siendo comparado después el número de casos reales y predichos en cada clase con el estadístico *chi-cuadrado* (χ^2) (Hair *et al.* 1999) (ecuación 35).

$$X^2 = \sum_{i=1}^{10} \frac{(O_i - E_i)^2}{E_i} + \sum_{i=1}^{10} \frac{(O_i^* - E_i^*)^2}{E_i^*} \quad (35)$$

Donde O_i son los valores observados dentro de cada grupo, E_i son los valores esperados dentro de cada grupo, O_i^* es la frecuencia de ocurrencia de ese grupo menos los valores observados y E_i^* es la frecuencia de ocurrencia de ese grupo menos los valores esperados.

En el modelo de regresión lineal, el coeficiente de determinación R^2 resume la proporción de varianza en la variable dependiente asociada con las variables predictoras, sin embargo, en el modelo de regresión logística, no es posible calcular un único estadístico R^2 , y hay que recurrir a otras pruebas estadísticas que indiquen la parte de la varianza de la variable dependiente explicada por el modelo. Hay dos R^2 (Pseudo- R^2) en la regresión logística y ambos son igualmente válidos: la R^2 de Cox y Snell, y la R^2 de Nagelkerke (Hair *et al.* 1999). Ambos Pseudo- R^2 devuelven un valor que oscila entre 0 y 1, de forma que se obtiene el valor máximo cuando el modelo asigna correctamente todos los casos en su respectiva categoría, y se aproxima a 0 cuando la asignación es deficiente.

Al igual que en el modelo de regresión lineal, la comprobación de la significación estadística del modelo de regresión logística se hace a nivel de modelo y a nivel de las variables independientes incluidas. A nivel de modelo se aplica un contraste *chi-cuadrado* (similar al estadístico F de la regresión lineal) para comprobar que el cambio del valor $-2LL$ (valor de verosimilitud final del modelo) respecto al valor de verosimilitud de referencia obtenido sin la inclusión de ninguna variable predictora en el modelo, es significativo ($p < 0,05$). Para los coeficientes estimados de las variables independientes y para la constante se utiliza el estadístico de Wald (similar al estadístico t de la regresión lineal, dándose por válido si su significación estadística se sitúa por debajo del nivel crítico ($p < 0,05$)).

También se puede comprobar la bondad del modelo de regresión evaluando la concordancia entre lo predicho y la ocurrencia-no ocurrencia real de la variable dependiente. Para ello se optó por dividir aleatoriamente la muestra compuesta por las parcelas de severidad en dos grupos: el 70% se utilizó para el ajuste y el 30% restante se reservó para la validación. Mediante una matriz de confusión o tabla de doble entrada en la que se comparan los hechos reales con los predichos, obteniéndose el porcentaje de acierto global del modelo (Hair *et al.* 1999) y el estadístico Kappa, se puede valorar la exactitud de la predicción. Tal como se comentó anteriormente, este índice delimita el grado de ajuste debido únicamente a la exactitud del modelo, desechando el causado por efectos aleatorios; tomando valores entre 0, acuerdo debido al azar, y 1, acuerdo perfecto (Pardo and Ruíz 2005, Chuvieco 2010). Por otra parte, las curvas de ROC (*Receiver Operating Characteristic*), constituyen una alternativa más para representar gráficamente la sensibilidad frente a la especificidad de un modelo o sistema clasificador binario según se varíe el umbral de discriminación. Mediante la comparación del área bajo la curva (AUC-*Area Under the Curve*) se puede cuantificar la capacidad discriminatoria del modelo. Esta área

posee un valor comprendido entre 0,5 y 1, donde 1 representa un valor diagnóstico perfecto y 0,5 nula capacidad discriminatoria (Hanley and McNeil 1982).

2.2.11 Representación cartográfica de las variables estimadas

Consiste en la estimación de las variables dependientes mediante la aplicación de los modelos obtenidos por regresión a la totalidad del área de estudio y su posterior representación en formato ráster. Para ello, en primer lugar hay que definir la resolución espacial. Las celdas deben tener una superficie similar a la de las parcelas empleadas en la fase de muestreo en campo. En segundo lugar, hay que obtener el valor de las variables independientes que intervienen en los modelos de regresión para cada una de las celdas utilizando un *software* específico, como FUSION 3.30 (McGaughey 2009). El comando “*GridMetrics*” genera un archivo “.csv” en el que figuran las variables LiDAR independientes para cada píxel del área de estudio y el comando “*CSV2Grid*” rasteriza dichas variables. Por último, solo resta aplicar los coeficientes de los modelos de regresión a cada celda para conseguir el valor de las variables dependientes para toda el área de estudio.

2.2.12 Clasificación digital basada en la combinación de datos

Tal como se exponía en el Capítulo 1, el último de los objetivos específicos de la tesis se centra en el desarrollo de una metodología para obtener una cartografía de tipos de combustible mediante la combinación de la información LiDAR-PNOA y una imagen óptica multiespectral del sensor SPOT 5. En este sentido, dicha perspectiva integradora ha permitido abordar una clasificación digital mucho más rica y completa de los tipos de combustible del área de estudio B, ya que se combinan criterios espectrales (propiedades ópticas de la vegetación) y estructurales (altura y densidad vegetal).

Sin embargo, contar con un amplio número de “bandas” de información (originales de la imagen multiespectral, índices de vegetación y derivadas de los retornos láser), acarrea ciertas complicaciones relacionadas con el procesamiento de la información y con la existencia de altas correlaciones entre ellas (redundancia). Para superar este inconveniente, se han utilizado dos técnicas de compactación de datos: el análisis de componentes principales (ACP) y la fracción de mínimo ruido (MNF-*Minimum Noise Fraction*), que permiten retener la mayor parte de la información contenida en todas las bandas originales empleadas en un menor número de nuevas bandas (Chuvieco 2010).

El ACP es una técnica estadística que tiene como objetivo sintetizar la información contenida en un grupo amplio de bandas en un nuevo conjunto más pequeño que conserva la parte más significativa de esa información. Es decir, los factores o componentes principales son “bandas-resumen” que no están correlacionados entre sí, ya que la información que contienen es específica, por lo que permiten extraer la estructura dominante de la escena. En términos generales, la obtención de los componentes principales de un conjunto de bandas puede resumirse en los siguientes pasos (Chuvieco 2010):

- A partir de la matriz de varianza-covarianza de las bandas originales se extraen los autovalores, que expresan la proporción de información original que retienen los nuevos componentes (longitud del componente). Este dato es fundamental para determinar qué componentes principales son más interesantes, teniendo en cuenta que el autovalor va decreciendo progresivamente desde el primer componente al último (habrá tantos componentes como bandas originales).
- El autovector es la ponderación que debe aplicarse a cada una de las bandas originales para obtener el nuevo componente principal.

- A partir de los autovectores se obtiene la matriz de correlación entre componentes principales y bandas de la imagen, lo que permite conocer el sentido espectral de estos componentes (dirección del componente).
- El último paso es obtener una imagen de los componentes principales a partir de la siguiente ecuación 36:

$$CP_j = \sum_{i=1,p} a_{i,j} ND_i + R_j \quad (36)$$

Donde CP_j indica el nivel digital ND correspondiente al componente principal j , $a_{i,j}$ es el coeficiente aplicado al ND de la banda i para generar el componente j y R_j es una constante que suele introducirse en cada componente para evitar valores negativos.

La técnica MNF consiste en aplicar dos ACP sucesivos. El primero permite eliminar la información redundante y re-escalar el ruido de los datos, empleando para ello una estimación de la matriz de covarianza del ruido. El segundo es un ACP convencional de los datos ya depurados, donde el ruido se estima a partir de los componentes con los autovalores más bajos y el análisis se centra en aquellos que retienen la mayor parte de la información original (Green *et al.* 1988).

Tras aplicar estas técnicas de reducción de la dimensionalidad de un conjunto de bandas diversas, el proceso de clasificación digital hace que la imagen multibanda (creada previamente) se convierta en otra imagen de idénticas dimensiones pero con la salvedad de que los ND de los píxeles ya no tienen relación con la radiancia detectada por el sensor, o con el estadístico que relaciona la altura y/o distribución de los retornos láser, sino que se trata de una “etiqueta” que identifica la categoría asignada a ese píxel (en este caso, tipos de combustible).

Básicamente, el esquema de la clasificación digital se compone de tres fases de acuerdo con Chuvieco (2010):

1. Entrenamiento: en esta fase se definen digitalmente cada una de las categorías que pretenden discriminarse, teniendo en cuenta su propia variabilidad en la zona de estudio. Este objetivo se logra seleccionando una muestra de píxeles de la imagen representativos de las categorías de interés, a partir de los cuales se calculan los ND medios y la variabilidad numérica de cada categoría en todas las bandas que intervienen en la clasificación.

De acuerdo a la forma en que son obtenidas las estadísticas de entrenamiento, los métodos de clasificación se dividen en dos grupos: supervisado y no supervisado. El método supervisado, que es el más habitual y el que se ha empleado en esta investigación, parte de un conocimiento previo del terreno, a partir del cual se seleccionan las muestras para cada una de las categorías o clases informacionales establecidas (tipos de combustible). Por el contrario, el método no supervisado procede a una búsqueda automática de grupos de valores homogéneos (clases espectrales) dentro del multibanda.

En cuanto al tamaño de la muestra de píxeles para el entrenamiento, se requiere seleccionar un mínimo de $m+1$ píxeles por categoría, siendo m el número de bandas que integran el análisis. No obstante, conviene superar ampliamente este límite mínimo. Respecto a la distribución, se debe abarcar las variaciones espaciales que se presentan en cada categoría en cuanto a su orientación, pendiente, densidad, vigor, contenido de humedad, etc. (Figura 21).

2. Asignación: en esta fase se agrupan los píxeles del multibanda en función de sus ND y se asignan a cada una de las clases previamente seleccionadas. Existen diferentes criterios para hacerlo, entre ellos el método de máxima probabilidad, que considera que los ND de cada clase se ajustan a una distribución normal. Esta función asemeja la distribución real de los ND en esa

categoría para poder calcular la probabilidad de que un píxel con un determinado ND sea miembro de ella. El cálculo se efectúa para todas las categorías que intervienen en la clasificación, asignando el píxel a aquella que maximice la función de probabilidad, es decir, con la que posea mayor probabilidad de pertenencia. Este clasificador paramétrico es uno de los más complejos y es el más empleado en teledetección, por su robustez y por ajustarse con más rigor a la disposición original de los datos.

3. Verificación: para comprobar la clasificación se utilizan los píxeles tomados en la fase de entrenamiento con objeto de extraer el valor de la imagen clasificada y obtener así una muestra de píxeles de verificación, que permitirá confrontar la clase “verdad-terreno” con la clase asignada por el clasificador. Con estos valores se puede generar una matriz de confusión que muestre los conflictos que se presentan entre categorías. En esta matriz cuadrada –de $n \times n$, donde n indica el número de categorías– las columnas suelen indicar las clases de referencia y las filas las categorías inferidas por la clasificación. La diagonal de esta matriz expresa el número de píxeles de verificación en donde se produce acuerdo entre la imagen clasificada y la realidad, mientras el resto de valores de las celdillas de la tabla hacen referencia a los errores de asignación. La relación entre el número de píxeles correctamente asignados y el total expresa la fiabilidad global de la clasificación. Los residuales en columnas indican tipos de cubierta real que no se incluyeron en la clasificación, mientras los residuales en filas implican cubiertas de la imagen clasificada que no se ajustan a la realidad. En definitiva, representan los errores de omisión y de comisión, respectivamente. De esta forma, no sólo se conoce la fiabilidad global de la clasificación, sino también la exactitud conseguida para cada una de las clases y los conflictos de asignación entre ellas.

También resulta de gran interés medir la diferencia entre el acuerdo imagen clasificada-realidad y el que cabría esperar simplemente por azar. Para este propósito, el índice Kappa, descrito anteriormente, resulta muy útil para comparar el acuerdo a nivel global y a nivel de clases.

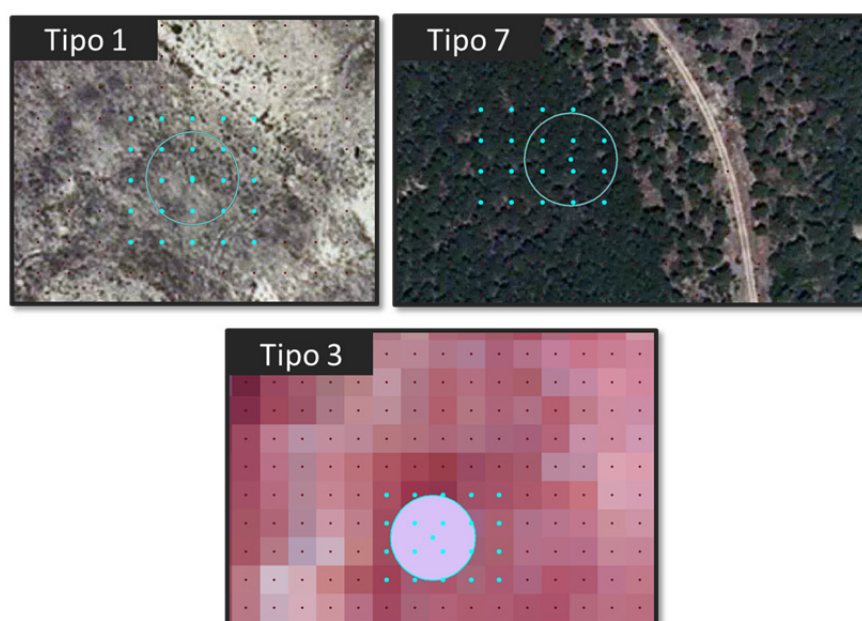


Figura 21. Ejemplo de selección de píxeles a partir de las parcelas de campo por tipo de combustible *Prometheus*. Como se puede apreciar, para la fase de entrenamiento de la clasificación se tomaron los píxeles considerados “verdad-terreno” del interior de las parcelas, así como los del entorno próximo, siempre y cuando tuviesen las mismas características.

3 APORTACIÓN DE LAS INVESTIGACIONES REALIZADAS

Los artículos y comunicaciones de congreso que conforman el cuerpo de la tesis doctoral se reproducen íntegramente a continuación, de acuerdo con la modalidad de tesis como compendio de publicaciones. Desde distintos puntos de vista, aportan un conocimiento útil sobre el manejo de los datos LiDAR-PNOA en el contexto del análisis de la estructura de la vegetación. Las dos primeras publicaciones se relacionan con el tratamiento adecuado de los datos (filtrado e interpolación) para mostrar cuáles son los procedimientos más precisos y conocer las fuentes de error. A continuación, las publicaciones restantes se centran en aplicaciones forestales concretas: la estimación de variables dasométricas o de inventario, la estimación del contenido de carbono almacenado en la biomasa arbórea, y la cartografía de la severidad post-fuego y de tipos de combustible forestal. Dichos trabajos responden secuencialmente a los objetivos específicos expuestos en el Capítulo 1 y se fundamentan en los datos y metodologías descritas en el capítulo anterior.

3.1 Comparación de métodos de filtrado para la clasificación de las nubes de puntos LiDAR del PNOA

RESUMEN

La tecnología LiDAR constituye un tipo de teledetección con potencial para ayudar en la cartografía, el monitoreo y la evaluación de los recursos forestales. A pesar de un creciente número de publicaciones revisadas por pares que documentan los métodos de filtrado de nubes de puntos LiDAR, parece haber poca información sobre la evaluación cualitativa y cuantitativa de éstos para seleccionar el más idóneo con la finalidad de crear MDE que permitan normalizar las alturas de los retornos para desarrollar aplicaciones forestales. Además, la mayoría de los algoritmos están implementados en *software* propietario, y unos pocos están disponibles abiertamente junto con información sobre su funcionamiento y sus resultados de aplicación.

Este artículo compara la precisión de siete métodos de filtrado de pulsos discretos, implementados en herramientas y *software* no propietario para la clasificación de las nubes de puntos proporcionadas por el Plan Nacional de Ortofotografía Aérea (PNOA). Se han seleccionado dos zonas de prueba con pendientes moderadas a fuertes y con varios tipos de cobertura de suelo. La precisión de cada algoritmo se ha evaluado utilizando 424 puntos clasificados manualmente.

El filtro MCC ha presentado el mejor rendimiento general, con un 83,3% de tasa de éxito y un índice Kappa de 0,67. Comparados con otros filtros, MCC y LAsTools equilibran bastante bien las tasas de error. Los matorrales rebrotadores, los troncos abandonados, los tocones y los restos leñosos que cubren el terreno, así como las pendientes topográficas superiores a 15° han sido los escenarios más problemáticos en el filtrado. La influencia de la densidad de puntos y del ángulo de exploración en el filtrado ha resultado menor, ya que los métodos morfológicos son menos sensibles a ellos.

A Comparison of Open-Source LiDAR Filtering Algorithms in a Mediterranean Forest Environment

Antonio Luis Montealegre, María Teresa Lamelas, and Juan de la Riva

Abstract—Light detection and ranging (LiDAR) is an emerging remote-sensing technology with potential to assist in mapping, monitoring, and assessment of forest resources. Despite a growing body of peer-reviewed literature documenting the filtering methods of LiDAR data, there seems to be little information about qualitative and quantitative assessment of filtering methods to select the most appropriate to create digital elevation models with the final objective of normalizing the point cloud in forestry applications. Furthermore, most algorithms are proprietary and have high purchase costs, while a few are openly available and supported by published results. This paper compares the accuracy of seven discrete return LiDAR filtering methods, implemented in nonproprietary tools and software in classification of the point clouds provided by the Spanish National Plan for Aerial Orthophotography (PNOA). Two test sites in moderate to steep slopes and various land cover types were selected. The classification accuracy of each algorithm was assessed using 424 points classified by hand and located in different terrain slopes, cover types, point cloud densities, and scan angles. MCC filter presented the best overall performance with an 83.3% of success rate and a Kappa index of 0.67. Compared to other filters, MCC and LAsTools balanced quite well the error rates. Sprouted scrub with abandoned logs, stumps, and woody debris and terrain slopes over 15° were the most problematic cover types in filtering. However, the influence of point density and scan-angle variables in filtering is lower, as morphological methods are less sensitive to them.

Index Terms—Airborne laser scanning, ground filtering algorithms, Mediterranean forest, open-source software.

I. INTRODUCTION

AIRBORNE light detection and ranging (LiDAR) has gradually become a common tool for collecting elevation information of surface targets with high precision and great density by calculating the time of flight taken for laser pulse travel between the LiDAR sensor and the target [1], [2]. Compared to the traditional photogrammetric method, the accuracies of the LiDAR measurements, approximately 0.15 m in altimetry and 1 m in planimetry under best conditions [3],

are unaffected by external light conditions, and its high spatial resolution outperforms the use of synthetic aperture radar (SAR) [4]. Furthermore, by distinguishing between the different reflections of a laser pulse, airborne LiDAR systems are capable of penetrating through vegetation, and recording the terrain beneath it [5]. Therefore, LiDAR has been widely used in digital elevation models (DEMs) generation, essential in environmental surveying and planning applications [6]. Since the raw LiDAR data contain a large number of points returned from various surface objects, such as buildings, bridges, electrical wires, and trees, these nonground/object points should be separated, the so-called LiDAR data filtering, prior to DEM construction. Conversely, bare-earth points need to be removed to accurately identify nonground objects [7].

According to several studies, the accuracy of a DEM developed with LiDAR data depends on: 1) the sensor and flight parameters, i.e., operating principles, scanner device, flight altitude, and speed [8], [9]; 2) the Earth's surface characteristics, i.e., topography and land cover [10]; and 3) the processing techniques used to create the DEM, i.e., filtering and interpolation methods, resolution, etc. [11]–[14]. However, Fisher and Tate [15] argue that relatively few studies have investigated error propagation between stages in DEM development. For instance, nonground points classified as ground, i.e., Type II or commission errors, may result in erroneous surface morphologies. Similarly, Type I or omission errors may lead to sparse ground points, failing to depict surface morphology [16].

On the other hand, a significant body of research has focused on LiDAR point classification, resulting in the development of several filtering methods, such as interpolation-based [5], [17]–[19], slope-based [20]–[22], segmentation-based [23], and morphological ones [24]–[26]. Meng *et al.* [27] identify some key assumptions in which most algorithms are based: 1) most of the terrain surfaces are locally autocorrelated and continuous, and the ground and nonground points exhibit an abrupt change in elevation; 2) as the terrain surface may be occluded, for instance by vegetation, the size of the local neighborhood should be adjusted to ensure that the terrain points are included; 3) the sizes of objects are within a limited range; 4) the lowest LiDAR points in a defined neighborhood have a higher probability of belonging to the terrain. Current approaches usually use these concepts independently or integrate several of them [28].

Filtering algorithms are typically tested using computer-simulated datasets for which the true ground is known [25]. In order to avoid the use of particular datasets and facilitate a meaningful comparison of performance between algorithms, Sithole and Vossleman [29] validated the performances of eight

Manuscript received November 03, 2014; revised March 29, 2015; accepted May 07, 2015. Date of publication July 06, 2015; date of current version September 12, 2015. This work was supported in part by the Government of Aragón (FPI Grant BOA 30, 11/02/2011) and in part by the Research Project of Centro Universitario de la Defensa de Zaragoza under Project 2013-04.

A. L. Montealegre and J. de la Riva are with the Department of Geography, University of Zaragoza, 50009 Zaragoza, Spain, and also with the GEOFOREST Research Group, Environmental Sciences Institute (IUCA), University of Zaragoza, 50009 Zaragoza, Spain (e-mail: monteale@unizar.es; delariva@unizar.es).

M. T. Lamelas is with the Centro Universitario de la Defensa de Zaragoza, 50090 Zaragoza, Spain, and also with the GEOFOREST Research Group, Environmental Sciences Institute (IUCA), University of Zaragoza, 50009 Zaragoza, Spain (e-mail: tlamelas@unizar.es).

Digital Object Identifier 10.1109/JSTARS.2015.2436974

1939-1404 © 2015 IEEE. Personal use is permitted, but republication/redistribution requires IEEE permission.

See http://www.ieee.org/publications_standards/publications/rights/index.html for more information.

classical filtering methods, in eight reference study sites (four urban and four rural), based on 15 samples representative of different environments, provided by the International Society for Photogrammetry and Remote Sensing (ISPRS) commission. They concluded that most filters perform well in flat and non-complex sceneries, but present problems in steep landscapes of dense vegetation. The last may be due to one of the assumptions of filtering algorithms: the bare-earth surface is smoother than the object's surface [14], [27]. Consequently, optimizing the algorithm parameters in large and topographically complex areas is still required [14], [18], [30]. Zhang and Whitman [30] pointed out the better performance of surface-based filters as more context information is used in the filtering process than in other strategies [31]. One of these methods presented by Axelsson [32] obtained better results in terms of total error in almost all reference study sites. From 2004 onward, several new filtering methods [26], [33]–[36] have been developed and evaluated based on the ISPRS dataset [37]. However, these methods perform even worse than that of Axelsson [32]. In addition, many reported results (e.g., [18] and [36]) correspond to researches applied to data with relatively high point densities and collected from low flight heights, typically 200–300 m above ground. In this sense, the need of filtering assessment arises when LiDAR datasets present medium–low nominal point density, as it is the case of the LiDAR data provided by the Spanish National Plan for Aerial Orthophotography (PNOA) with 0.5 points/m².

Most filtering methods offer a strong theoretical background, but they are still application specific, as they require additional information about the studied area to achieve satisfactory results. For example, knowledge-based methods, specially developed for characterizing cityscapes, have been exploited to include terrain information [38], [39] but extensive databases, sometimes difficult to obtain, are required [40]. Recently, statistically based methods have been introduced [36], [40]–[42] to achieve parameter-free methods based on skewness balancing. However, a set of conditions needs to be satisfied, e.g., a minimum number of ground points to be used [40]. Consequently, they are unable to remove attached objects and preserve ground points with irregular height distribution [36].

Despite the development of new methods and the widespread use of LiDAR-derived DEMs, filtering has been proven to be exceptionally difficult to automate especially in applications with large datasets in areas of diverse terrain characteristics [33], [36], [43]. Furthermore, there is little guidance in the literature regarding the selection of parameters, e.g., thresholds and window sizes, to be included to optimize filtering [6], [19], [30]. In fact, point classification algorithms commonly applied by LiDAR vendors are proprietary knowledge, being very often gray- or black-box approaches, not readily available for independent validation and comparison. Fortunately, in recent years open-source algorithms designed for discrete-return LiDAR data have been developed, which can be independently tested, evaluated, and compared [44], [45].

Due to the lack of an optimal filtering algorithm, a quality control becomes necessary to select the most suitable in a particular context. The influence of different variables like pulse density, terrain slope, and vegetation on the vertical accuracy

of LiDAR-derived DEMs has been commonly assessed [12]; however, little research has focused on the comparison of different point classification algorithms. Therefore, the research objectives of this paper are: 1) to evaluate the relative performance of seven different well-known filtering methods available in nonproprietary software, the progressive TIN densification algorithm (LASTools), the weighted linear least-squares interpolation-based method (FUSION), the multiscale curvature classification (MCC), the interpolation-based filter (BCAL), the elevation threshold with expand window method (ETEW-ALDPAT), the progressive morphological filter (PM-ALDPAT), and the maximum local slope algorithm (MLS-ALDPAT), in medium–low density point clouds captured in a forest environment; 2) to determine the influence of terrain slope, land cover, point density, and scan angle in the filtering error; and 3) to provide guidance for users of PNOA LiDAR point clouds to select the most suitable filtering algorithm to be applied in a Mediterranean pine forest using nonproprietary tools.

II. MATERIAL AND METHODS

A. Study Area

The study area consists of two sample sites, T1 (2 km × 2 km) and T2 (4 km × 2 km), located in the central Ebro valley (41° 56' N, 0° 56' W), sited northeastern Spain (Fig. 1). The Ebro Basin constitutes the northernmost semi-arid region in Europe and stretches from the Pyrenees range, in the north, to the Iberian range, in the south.

This area presents a Mediterranean climate with continental features. Annual precipitation is low, averaging 350 mm, and presents an irregular distribution during the year, mostly concentrating in autumn and spring. Moreover, the study area is characterized by cold winters, with monthly mean temperature about 7 °C, and hot, dry summers, with temperatures about 24 °C [46]. With respect to topography, the area presents a hilly relief, with elevation ranging from 400 to 750 m above sea level, and moderate-to-steep slopes (Fig. 2).

In the two selected sites, Aleppo pine forests (*Pinus halepensis* Mill.) cover 528 ha and pine terrace plantation 30 ha, being interspersed with evergreen shrubs, dominated by *Quercus coccifera* L., *Juniperus oxycedrus* L. subsp. *macrocarpa* (Sibth. & Sm.) Ball and *Thymus vulgaris* L. covering a total of 302 ha, and cereal crops account for 115 ha (see Table I).

The forest presents a homogeneous structure, an average canopy height of 6.5 m and an average biomass of 45 t/ha. Old stands reach 12–13 m in height and 90 t/ha of biomass [47]. In addition, in the last century, the study site has been recurrently affected by fire, some areas being burned even twice. Two scars of wildfires developed in June 1995 and August 2008, which consumed 5300 ha of forest, are distinguishable nowadays [48]. Particularly, the west end of T2 is covered by 232 ha of coniferous forest affected by a wildfire in 2008. Currently, the vegetation of this area is dominated by shrub species that colonize rapidly, while the succession to forest needs considerably longer time [49]. Thus, this study area is characteristic of a Mediterranean environment, repeatedly affected by wildfires

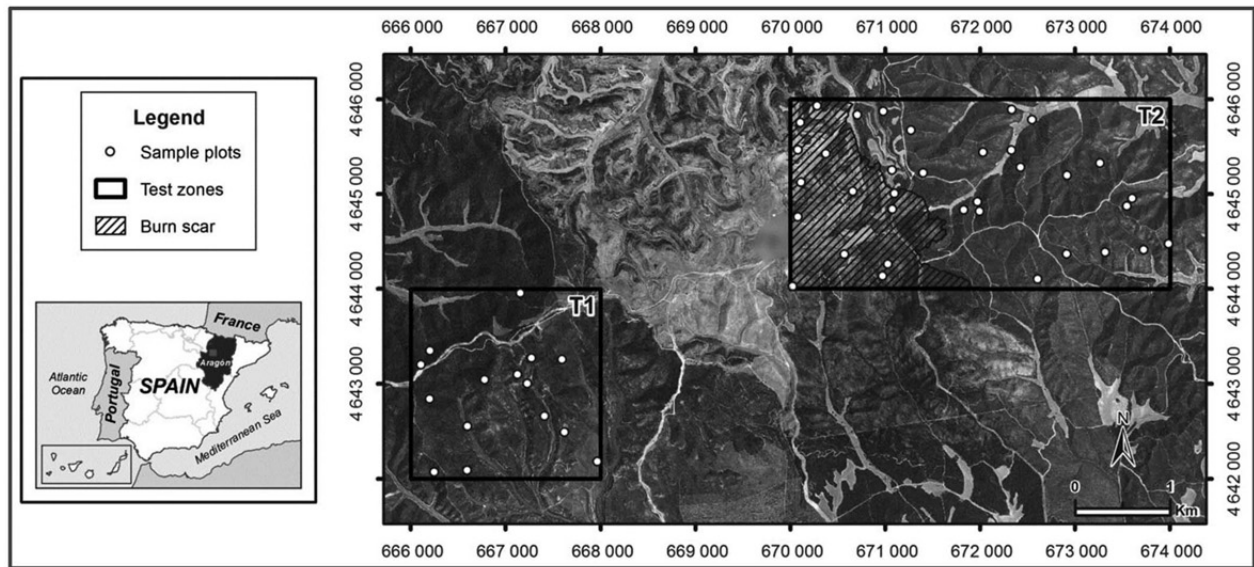


Fig. 1. Study area (T1 and T2 sites) and the 50 random sample plots. As background a high spatial resolution orthophotography (source: PNOA 2009).

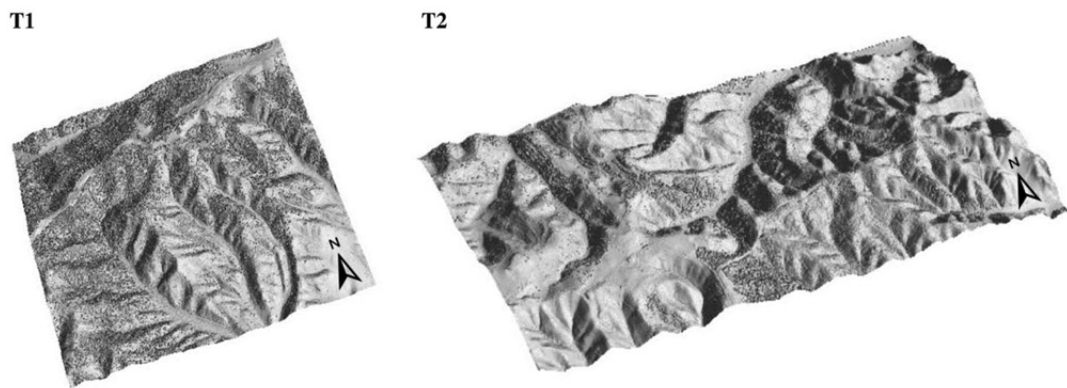


Fig. 2. 3-D shaded surface models from unfiltered LiDAR PNOA point clouds of test sites: T1 and T2.

TABLE I
SUMMARY OF TEST SITES T1 AND T2 CHARACTERISTICS

Test site	Mean elevation (m)	Elevation range (m)	Mean slope (degrees)	Slope range (degrees)	Coniferous forest (%)	Pine terrace plantation (%)	Shrub (%)	Crops and grasslands (%)	Burned area (%)
T1	661	162	16	38	27	8	51	14	0
T2	627	262	16	41	52	0	12	7	29

[50]. As a result, the landscape is a patchwork of bare ground, fields, shrubs, tree skeletons, and stands of coniferous forest.

B. LiDAR Data Acquisition

The LiDAR data were provided by the PNOA (http://www.ign.es/PNOA/vuelo_lidar.html) and captured in several surveys conducted between January 22 and February 5, 2011, using an airborne Leica ALS60 discrete return sensor. Data were delivered in three 2 km × 2 km tiles of raw data points in LAS binary file, format v. 1.1, containing x- and y-coordinates (UTM Zone 30 ETRS 1989), ellipsoidal elevation z (ETRS

1989), with up to four returns measured per pulse and intensity values from a 1064-nm wavelength laser. The resulting LiDAR point density of test areas was 1 point/m² with a vertical accuracy higher than 0.20 m. The properties of the LiDAR acquisition are summarized in Table II. It should be noted that all returns were used in the processing.

C. Software and Ground Filtering Algorithms Evaluated

The filters used to separate ground and nonground point measurements in the selected test areas include: progressive TIN densification (LASTools), weighted linear least squares

TABLE II
LiDAR DATA SPECIFICATIONS AND ACQUISITION PROPERTIES

Property	Value
Sensor	ALS60
Scanning method	Oscillating plane mirror (saw-tooth pattern)
Date	January and February 2011
Mean flying height above ground (m)	3012
Mean flying speed (km/h)	241
Laser return density (point/m ²)	1.0
Field of view (degrees)	40
Beam divergence angle (mrad)	0.22
Scan angle (degrees)	±22°
Total extension of LAS files (km ²)	12
Point count	18 495 618
Altimetric accuracy (RMSE in m)	0.20

interpolation-based (FUSION), multiscale curvature classification (MCC), interpolation-based (BCAL), elevation threshold with expand window (ETEW), progressive morphological (PM), and maximum local slope (MLS). All filters are surface-based methods, as the core step of this kind of methods is to create a parametric surface approximating the bare earth with a buffer zone that defines a region in three-dimensional (3-D) space where ground points are expected to reside [29]. Depending on the way of creating the surface, these methods can be further divided into interpolation-based, progressive TIN densification, and morphology-based subcategories [31]. An overview of the software and a characterization of the filters associated with them is given in Table III and described below in more detail.

1) *LAStools*: *LAStools* is a suite of LiDAR data-processing tools programmed by Martin Isenburg. The tool *lasground* was used to label each point as ground point or not (<http://rapidlasso.com/lastools/>). This tool implements the method proposed by Axelsson [32], [51], which is based on a grid simplification. First, this algorithm divides the whole-point dataset into tiles and selects the lowest points in each tile as the initial ground points. Then, a triangular irregular network (TIN) of those ground points is constructed as the reference surface. In each triangle of the TIN, one of the unclassified points is added to the set of ground points following two criteria: the point's distance to the TIN facet and the angle between the TIN facet and the line connecting the point with the closest vertex of the facet must not exceed a given thresholds. Before the next iteration, all ground points classified in the current iteration are added to the TIN. In this way, the triangulation is progressively densified until all points are classified as either ground or object [6], [31]. In practice, parameterization of *lasground* consists of the selection of two settings: the terrain type (a step size of 5 m, suitable for forest and mountains, was selected) and the granularity, i.e., how much computational effort to invest into finding the initial ground estimate (the options "default" and "fine" were selected).

2) *FUSION*: *FUSION* v. 3.30 software [52] was developed at the U.S. Forest Service Pacific Northwest Research Station (<http://forsys.cfr.washington.edu/fusion/fusionlatest.html>). The command *groundfilter* used to generate a bare-earth surface is adapted from Kraus and Pfeifer [5] and is based on linear prediction [53], which belongs to the category of so-called interpolation-based filters. These type of filters usually fit a surface to the data and iteratively classify points based on a function to assign weights to each point (p_i) based on its residual (v_i) from the fitted surface. In the first iteration, all points are given equal weights and an averaging surface model is computed, so the residuals of the data points relative to the surface are calculated [54]. If the measured points lie above it, they have less influence on the shape of the surface in the next iteration, and vice versa [5], i.e., ground points are more likely to have negative residuals, so they are given more weight in subsequent iterations and thus they attract the computed surface toward themselves [6], [31], [54]. The *groundfilter* command computes the weights for each LiDAR point using (1)

$$p_i = \begin{cases} 1 & v_i \leq g \\ \frac{1}{1+(a(v_i-g)^b)} & g < v_i \leq g+w \\ 0 & g+w < v_i \end{cases} \quad (1)$$

The parameters a and b determine the steepness of the weight function. The *FUSION* manual recommends values of 1.0 and 4.0 for a and b , respectively, in most applications. The shift value g determines which points are assigned a maximum weight of 1.0. Points located a higher distance than g below the surface are assigned a weight of 1.0. The above-ground offset parameter w is used to establish an upper limit to points having an influence on the intermediate surface. Points above the level defined by $(g+w)$ are assigned a weight of 0.0. In the current implementation, values for g and w are fixed throughout the filtering run. Kraus and Pfeifer [5] used an adaptive process to modify the g parameter for each iteration. After the final iteration, default is 5, ground points are selected using the final intermediate surface. All points with elevations that satisfy the first two conditions of the weight function are considered bare-earth points [52]. In the absence of guidance to select appropriate values and given the numerous possible combinations, experimentation by setting different parameter values was performed.

3) *Multiscale Curvature Classification*: *MCC-LiDAR* v.2.1 is an open-source command-line tool developed to process discrete-return LiDAR data in forest environments and is available on <http://sourceforge.net/p/mcclidar/wiki/Home/>. It classifies data points as ground or nonground using the MCC algorithm, developed by Evans and Hudak [19] at the Moscow Forestry Sciences Laboratory of the USFS Rocky Mountain Research Station.

Like *FUSION* 3.30 software [52], *MCC* is an iterative-interpolation-based filter. The *MCC* algorithm operates by discarding returns that exceed a threshold curvature, calculated from a surface interpolated using a thin-plated spline. Through three successively larger scale domains that define the processing window size, the algorithm iterates until the number of remaining returns changes by less than 1%, less than 0.1%, and

TABLE III
EVALUATED FILTERING ALGORITHMS AND KEY PARAMETERS

Software/tool	Developer	Filter description	Key parameters
LASTools—"lasground.exe"	Martin Isenburg adapted from Axelsson [32], [51]	Progressive TIN densification	Terrain type, granularity
FUSION—"groundfilter.exe"	Robert J. McGaughey adapted from Kraus and Pfeifer [5]	Interpolation-based	Weight assigned to points (g, w, a, b parameters), cell size
MCC—LiDAR	Evans and Hudak [19]	Interpolation-based	Scale, curvature
BCAL LiDAR Tools—"Perform Height Filtering"	Streutker and Glenn [55]	Interpolation-based	Return number, canopy spacing, threshold value, interpolation method, maximum iteration
ALDPAT—"Elevation Threshold with Expand Window (ETEW) filter"	Zhang and Whitman [30]	Morphology-based	Initial cell size, slope factor, number of iteration
ALDPAT—"Progressive Morphological (PM) filter"	Zhang <i>et al.</i> [25]	Morphology-based	Initial cell size, window size, slope, initial threshold
ALDPAT—"Maximum Local Slope (MLS) filter"	Vosselman [20]	Morphology-based	Cell size, search radius, slope threshold

finally less than 0.01% in the three scale domains, respectively [19], [44].

There are two parameters that must be defined in the command-line syntax to run MCC: the scale parameter (s) and the curvature threshold (t). The optimal scale parameter is a function of the scale or size of the objects and point spacing of the LiDAR data. Since the point spacing of the test areas is 1 m, a scale parameter of 1 was determined and three values were tested for the curvature threshold, 0.3, 0.4, and 0.5 as recommended by the developers of this algorithm.

4) *BCAL LiDAR Tools*: BCAL LiDAR Tools v.1.5.1 was originally developed by David Streutker from the BCAL of Idaho State University and is distributed through <http://bcal.boisestate.edu/tools/lidar/>. BCAL LiDAR Tools have also been used by ITT Exelis to develop their own proprietary LiDAR extension for ENVI software. The *Perform Height Filtering* tool designed for optimal performance in shrub–steppe ecosystems [55] classifies LiDAR data into ground and vegetation. It is a grid-based classification algorithm that first identifies the lowest elevation point in a search area determined by the user, and then creates a surface by interpolating these lowest points [55]. This tool uses several interpolation methods, such as cubic spline, inverse distance, inverse multiquadric, linear, natural and nearest neighbor, polynomial regression, and thin plate spline. In subsequent iterations, any point that lies on or below the previous iteration's surface is classified as ground and is included in subsequent iterations until no unclassified returns remain below the interpolated surface. Then, all unclassified returns above the surface are classified as object [55].

5) *Airborne LiDAR Data Processing and Analysis Tools*: ALDPAT v.1.0 was developed by the National Center for Airborne Laser Mapping (NCALM). This software implements several algorithms to classify the ground and nonground LiDAR measurements [56] and is available on <http://lidar.ihr.fiu.edu/index.html>. This group of morphology-based filters derived from mathematical morphology [57] is based on the combination of two basic operations, the so-called closing (erosion after dilation) and opening (dilation after erosion), to

determine the minimum and maximum points within a certain structure element (window) and to remove the object returns [6], [24]. The collection of algorithms used is described below.

a) *ETEW filter*: This filter is based on the Zhang and Whitman [30] algorithm. The LiDAR dataset is subdivided into an array of square cells and all points, except the minimum elevation, are discarded. In the next iteration, the cells are increased in size and the minimum elevation in each cell is determined. Then, all points with elevation higher than a threshold above the minimum are discarded. This process is repeated for increasing cells and thresholds in size until no points from the previous iteration are discarded [30], [56].

b) *PM filter*: Zhang *et al.* [25] developed a PM filter to differentiate between ground and nonground points based on elevation differences between cells in a moving window using morphological operations such as openings and closings. The PM filter removes the measurements in different sized nonground objects, while preserving ground data to derive a parametric surface model, by gradually increasing the window size and using elevation thresholds. The process ends when the size of the filtering window is larger than the predefined maximum size of nonground objects. The cell size is usually selected to be smaller than the average spacing between LiDAR measurements to preserve the highest amount of points. If no measurement exists in a cell, the value of its nearest neighbor is assigned.

c) *MLS filter*: Vosselman [20] developed a filter that describes the maximum admissible height difference within a structure element (circular window) as a function of the distance calculated as the horizontal euclidean distance between two points, the so-called local slope concept. Since terrain slope is usually different from the slope observed between the ground and the tops of trees, this slope or gradient difference can be used to separate ground and nonground measurements from a LiDAR dataset. Each point measurement from the cloud is assigned into a cell of an array in terms of its x - and y -coordinates. If more than one point falls in the same cell, the one with the lowest elevation is selected. A point is classified

as bare earth if the maximum value of slope between this point and any other point (height difference) within a given radius (distance) is less than a predefined threshold. In practical applications, the predefined radius for the structure element is typically set to 5 m. Parameters of the filter function were determined based on the maximum terrain slope found in the area and from the height precision of the laser points [6].

D. Reference Data for Validation

There are two basic errors in filtering LiDAR data. The first one is to classify nonground measurements as ground points (Type II error), and the second one (Type I error) is to select ground points as nonground measurements [29]. Since all filtering methods are subject to these two errors, results should be examined. The validation of the point classification with the whole-point cloud is impractical due to the large number of measurements recorded. An alternative approach proposed by Zhang and Whitman [30] and Zhang *et al.* [25] that examines a sample of randomly selected test points was selected. First, a set of 50 x - and y -coordinates was selected randomly within the bounds of the T1 and T2 test sites. The sampling protocol ensured that these locations covered the terrain and vegetation variability of the study area (Fig. 1). Then, LiDAR measurements that fell within 3 m of distance to the x - and y -coordinates were selected as test points. Finally, the 424 test points selected were classified manually using a high spatial resolution orthophotography provided by the PNOA mission, the intensity image created from the LiDAR data, as well as the 3-D visualization of the points. In the case of points with higher difficulty to be classified, a field campaign was conducted. They were staked out in field using a Leica VIVA GS15 CS10 GNSS real-time kinematic (RTK) global positioning system to confirm the manual filtering and to be as precise as possible.

E. Validation

The evaluation approach used by Sithole and Vosselman [29] is adopted here to assess quantitatively and qualitatively the performance of the seven filtering methods applied. Sithole and Vosselman [29] proposed three accuracy metrics to quantitatively analyze the performance of a filter: 1) Type I error—rejection of bare-earth points (2); Type II error—acceptance of object points as bare earth (3); and total error (4)

$$Type\ I_{Error} = \frac{a}{BE} \quad (2)$$

$$Type\ II_{Error} = \frac{b}{OBJ} \quad (3)$$

$$Total_{Error} = \frac{a+b}{BE+OBJ} \quad (4)$$

where a is the number of ground points incorrectly identified as object, b is the number of object points incorrectly identified as ground, and BE and OBJ are the total number of bare earth and object points in the reference data, respectively. The total error rate is equal to the sum of all wrong classifications divided by the total number of points in the dataset [37]. Alternatively,

TABLE IV
QUALITATIVE COMPARISON OF FILTERS

Assessed objects	Filters				
	MCC	LAS tools	BCAL	ALDPAT	FUSION
<i>Discontinuity</i>					
Preservation (steep slopes)	***	***	*	*	**
Sharp edges (roads, terracing)	***	***	**	**	**
Small reliefs	***	***	*	*	**
<i>Vegetation</i>					
Vegetation	***	***	***	**	***
Vegetation on slopes	**	**	**	**	**
Low vegetation	**	**	**	**	**

*Poor (item not filtered most of the time, i.e., <50%).

**Fair (item not filtered a few times).

***Good (item filtered most of the time, i.e., >90%).

success rate was calculated as the ratio between points correctly classified and the total number of points in the dataset.

Recently published algorithms [7], [33], [37] utilize Cohen's Kappa index [58] as a measure of accuracy as this index seems to be a more robust measurement than a simple percentage [59]. This statistical coefficient measures the overall inter-rater agreement, accounting for the possibility of chance in the observed frequencies. Kappa index ranges generally from 0 to 1, although negative numbers are possible. According to Landis and Koch [60], values of Kappa below 0.40 present poor agreement, between 0.40 and 0.75 are considered as good agreement, and above 0.75 generally reflect excellent agreement.

Furthermore, the effect in the filtering error of four variables, such as terrain slope, land cover type, point density, and scan angle, was examined. In this sense, the error metrics were stratified based on several categories established in these variables.

Finally, the qualitative assessment of all filter procedures consists of a visual examination and comparison of a shaded relief of the DEMs developed with the filtered datasets.

III. RESULTS

A. Qualitative Assessment

The results of the qualitative assessment of filters performed following the same criteria proposed by Sithole and Vosselman [29] are summarized in Table IV. Figs. 3 and 4 exemplify miscellaneous difficulties in filtering in two samples representative of the study area. First, Fig. 3 corresponds to a complex mixture of Aleppo pine forest on a steep slope (i), pine terrace plantations (ii), and shrub vegetation about 2 m height (iii). Second, Fig. 4 consists of the area affected by fire in 2008 where a high number of bare-earth points exist (i), low vegetation on slope (ii), and pine forest (iii). The visualizations of the filtered shaded reliefs exhibit nearly the same appearance in all the samples, i.e., most of nonground object points were removed by the five filters. However, several Type I and Type II errors were committed by the filters.

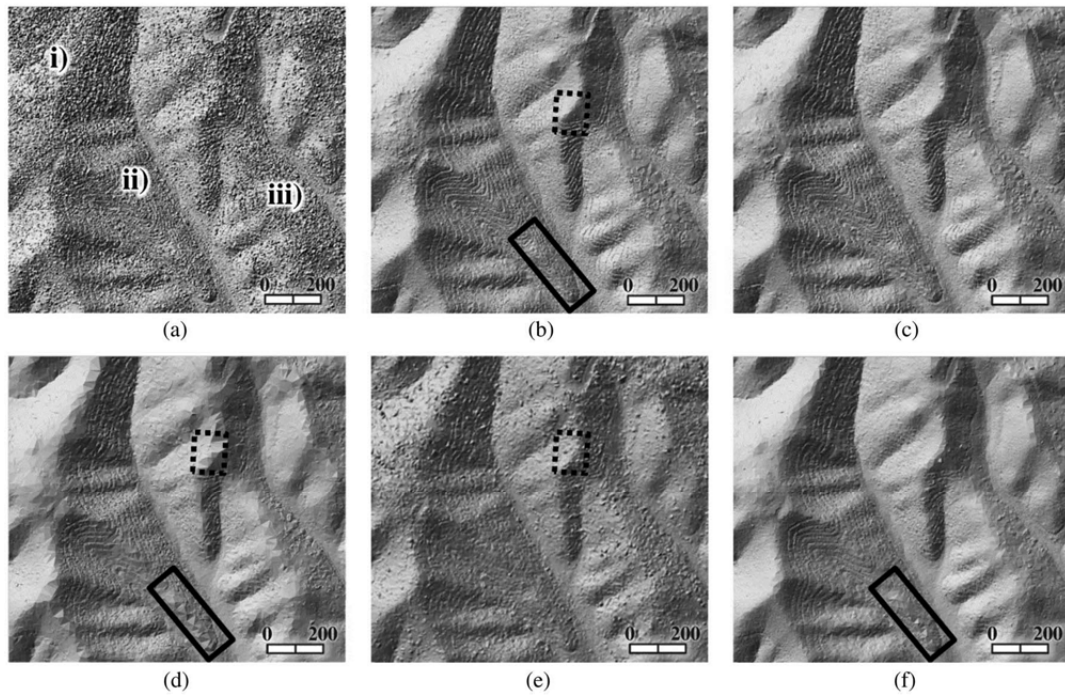


Fig. 3. 3-D surface models generated from the (a) unfiltered data and the filtered bare-earth points using (b) MCC, (c) LASTools, (d) BCAL, (e) ALDPAT-PM, and (f) FUSION filtering methods. The sampling area is characterized by (i) Aleppo pine forests on a steep slope, (ii) pine terrace plantations, and (iii) shrub vegetation about 2-m height.

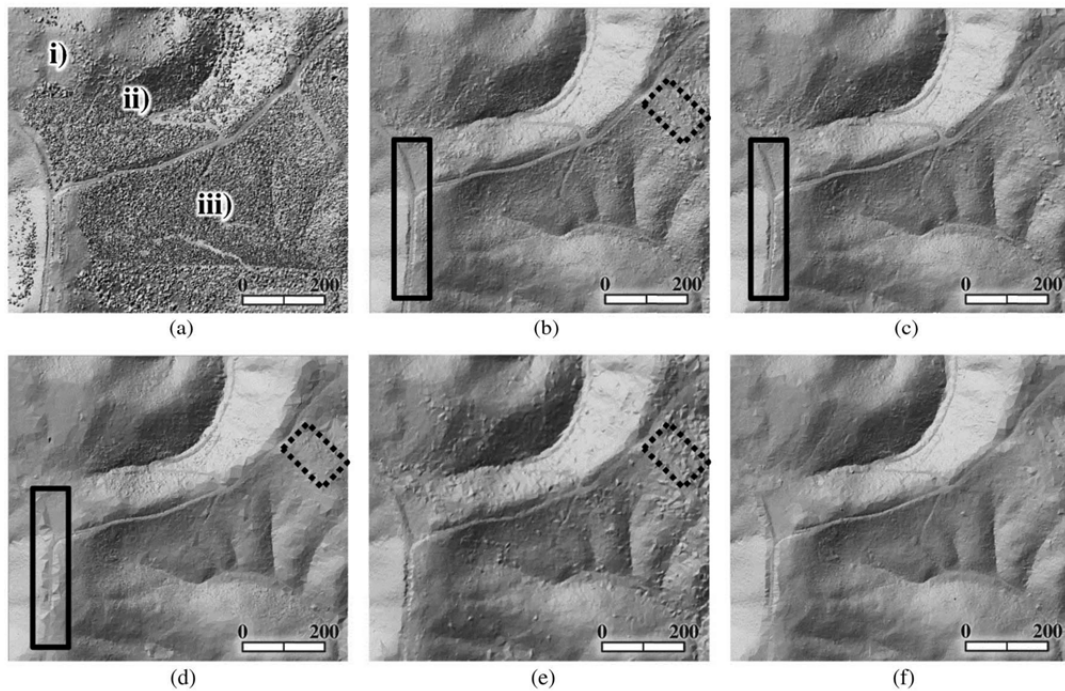


Fig. 4. 3-D surface models generated from the (a) unfiltered data and the filtered bare-earth points using: (b) MCC, (c) LASTools, (d) BCAL, (e) ALDPAT-PM, and (f) FUSION filtering methods. The sampling area is characteristic of an area affected by (i) fire in 2008, (ii) low vegetation on slope, and (iii) pine forest.

For example, a small mound (dashed line rectangle in Fig. 3) was removed completely by the BCAL filter [Fig. 3(d)] or partially contaminated by the shrub returns in ALDPAT-PM filter

[Fig. 3(e)], since vegetation points were mistakenly classified as ground points. On the other hand, when the bare earth is piecewise continuous, some filters will operate as they would

TABLE V
TYPE I AND TYPE II ERRORS, SUCCESS RATE, AND COHEN'S KAPPA INDEX (p -value ≤ 0.05) OF DIFFERENT FILTERING METHODS

Filtering method	Optimum parameters	Type I error (%)	Type II error (%)	Success rate (%)	Kappa index
MCC	Scale 1, curvature 0.3	12.7	20.8	83.3	0.67
LAStools	Terrain type forest and hills, granularity fine	20.8	13.7	82.8	0.66
BCAL	All return number, canopy spacing 5, threshold value 0, inverse distance-first order interpolation method, maximum iteration 15	38.2	7.5	77.1	0.54
ALDPAT-PM	Initial cell size 1, window size 3, slope 1.2, initial threshold 0.20	39.6	14.6	72.9	0.46
FUSION	$g=2.5$, $w=2$, $a=1$, $b=4$, cell size 2	61.3	2.8	67.9	0.36
ALDPAT-ETEW	Initial cell size 1, slope 1, number of iteration 5	35.8	31.6	66.3	0.33
ALDPAT-MLS	Cell size 1, search radius 15, slope threshold 0.3	75.0	0.0	62.5	0.25

Only the results of the best settings are shown for each type of filter along with the parameters implemented.

on objects, i.e., Type I errors [29]. This is precisely what is observed in continuous line rectangles in Fig. 3(d) and (f), compared to the continuous line rectangle in Fig. 3(b), where discontinuities in the bare earth are lost, i.e., the pine terrace plantations are moderately preserved.

The roughness along the slope, denoted by the dashed line rectangles in Fig. 4(b) and (e), is created by scattered Type II errors in very low objects, i.e., presence of shrub vegetation dominated by *Quercus coccifera*. Compared to the dashed line rectangle in Fig. 4(d), the surface is generally smoother due to the high percentage of Type I errors committed by the BCAL method, as ground points were classified as nonground ones.

Finally, MCC and LAStools preserved quite well the limits of roads and forest tracks [see continuous line rectangles in Fig. 4(b) and (c)], but the rest of filtering methods present certain problems with these features. For example, the sharp edge of the road was removed partly by the BCAL, ALDPAT-PM, and FUSION filters. These algorithms tend to dilate the boundaries of areas with lower elevation relative to their neighbors [30], leading to a distortion of the surface because of this Type I error. An example of this "over filtering" can also be observed in Fig. 3(d).

B. Quantitative Assessment

First of all, to achieve optimal results, several combinations of filter parameters were applied. As a result, the parameter set presenting the minimum total error was chosen as the optimum following Hu *et al.* [59]. In Table V, the computed errors can be observed. Type I, Type II, and total errors ranged from 12.7% to 75.0%, from 0.0% to 31.6%, and from 16.7% to 37.5%, respectively. MCC filter presented the lowest overall error (16.7%), while ALDPAT-MLS and ALDPAT-ETEW filters achieved the highest overall error, 37.5% and 33.7%, respectively. Type II error in MCC was 20.8%, in comparison with the 12.7% of Type I error. The inclination to commit Type II errors may not be a handicap for this filtering method, taking into consideration that Type II errors can be more easily handled by manual editing than Type I errors [29], [31], [59]. In this sense, the results obtained in MCC filtering indicated no severe biases in the

classification, as this method appropriately balanced Type II and Type I errors [19].

The second filter with the best percentage of success was LAStools, which differs from MCC in the distribution of errors. In this case, only a 13.7% of object points were misclassified as bare earth, while 20.8% of points were rejected as bare earth. A possible reason is that although the surface represented by the TIN is able to handle point density variations, the local ground surface is only relevant with three vertices and is expressed by a triangle that is simply a plane, being quite sensitive to noise [59].

With respect to ALDPAT-MLS method, it presented a high percentage of Type I errors, 75.0%, but it did not obtain Type II errors. This algorithm tends to discard many ground points, thus keeping a more sparse set of ground measurements to generate the surface. It should be noted that the sparseness of these points does not necessarily lead to a worse DEM interpolation [30]. Compared to the interpolated-based methods, the sensitivity to parameters is the major drawback in ALDPAT morphological-based filters [36], [59]. Nevertheless, the ALDPAT-PM filter committed fewer errors than ETEW and MLS methods.

Finally, to analyze the effects of the terrain slope, cover type, point density, and scan-angle variables on error rates, four comparative experiments, which are depicted in Figs. 5–8, were performed.

1) *Terrain Slope*: The influence of the terrain slope was evaluated using two categories: smooth slopes ranging from 0° up to 15° , and steep slopes higher than 15° (Fig. 5).

Discontinuities or height differences are a key assumption to separate the bare earth from objects. Consequently, points significantly above their neighbors are assumed to belong to objects, but this assumption becomes difficult when the slope of the terrain increases [29]. Therefore, as expected, most filters had difficulties on steep slope higher than 15° (e.g., Fig. 3), increasing their error rates with the terrain complexity. In this sense, all analyzed filters, except ALDPAT-MLS, increased their Type I error. Even the filters with the lowest total error (20%) in slopes higher than 15° , MCC and LAStools filters increased their error in this complex environments. Filters based on slope, such as ALDPAT-MLS may

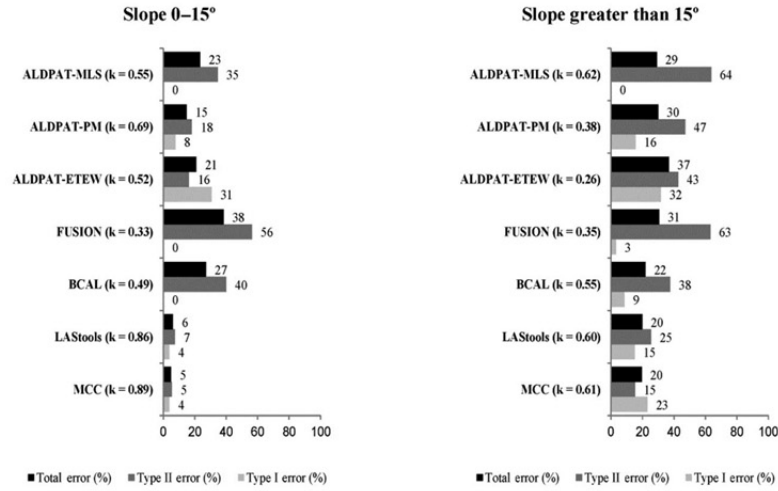


Fig. 5. Effect of terrain slope on filter performance and Cohen's Kappa index (k) (p -value ≤ 0.05).

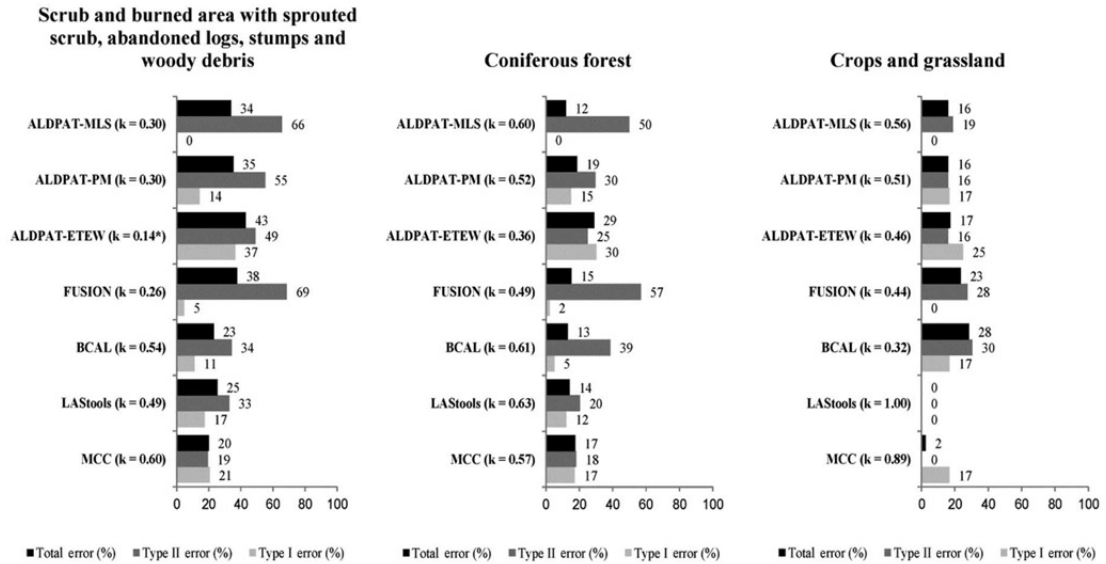


Fig. 6. Effect of cover types on filter performance and Cohen's Kappa index (k) (p -value ≤ 0.05). (*) Nonsignificant Kappa value (p -value > 0.05).

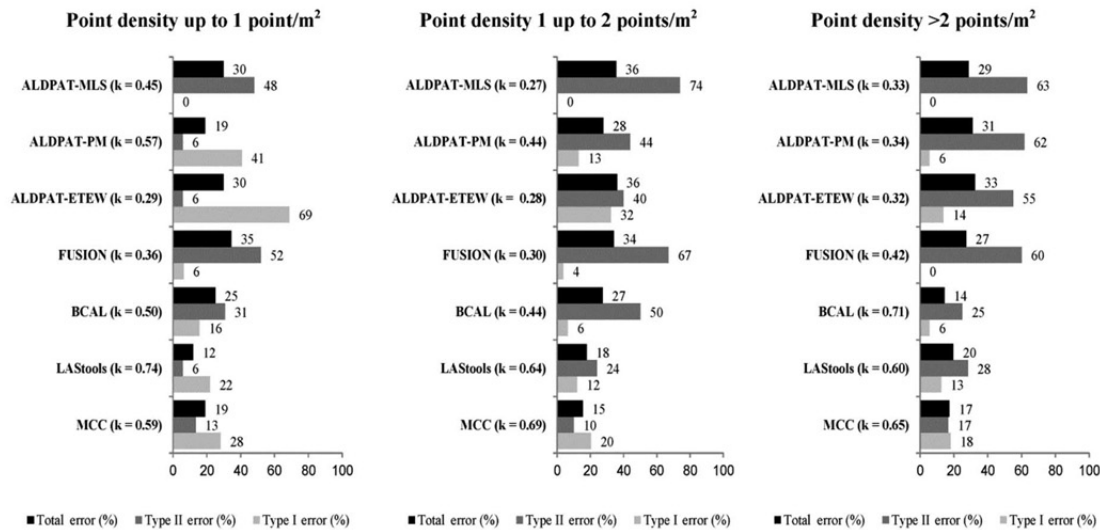
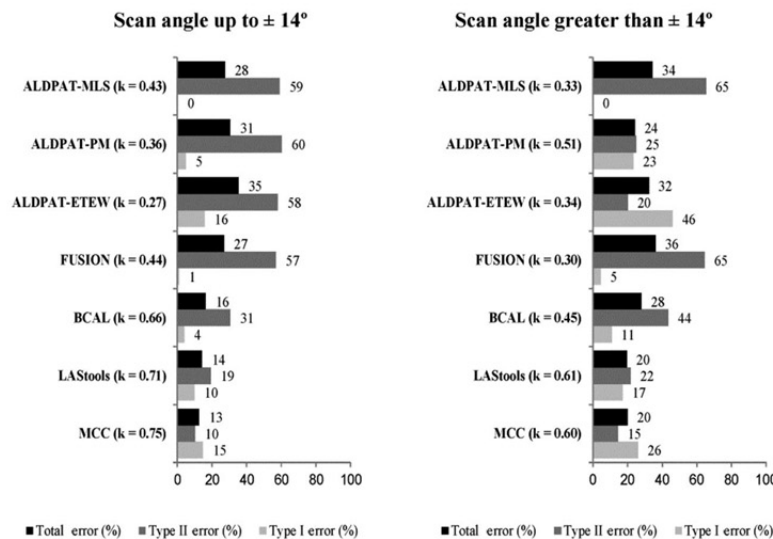
mislabel points as nonground in areas with slopes larger than the maximum ground slope threshold. In this regard, the study area, which presents a variety of situations, from even areas to very steep ones, is challenging for ground filters because of the difficulty of selecting appropriate slope and elevation thresholds [27]. Type II errors also increased with the change in slope, mainly in the case of morphology-based filters. Furthermore, the ALDPAT filters had the highest difficulty in bare-earth discontinuities preservation, as was explained above (Fig. 3).

2) *Cover Types*: We also compared the performance of the algorithms within each of the main cover types of the test areas. As can be observed in Fig. 6, MCC filter was significantly better than the other six methods in terms of total error (20%) and Type II error (19%), with a good level of agreement (Kappa 0.60), particularly in scrubs and in burned areas occupied with sprouted scrub and abandoned logs. As expected, Type II errors

were higher in both types of cover due to the presence of attached and low objects such as abandoned logs, stumps, and small seedlings, as well as shrubs, which produce the classification of objects as bare earth. Points returned from shrub cover are commonly mislabeled as ground surface, as in relatively steep terrain, the slope and elevation difference between the shrub and neighboring ground points are similar to those between ground points and neighboring ground points [27].

In coniferous forest cover, ALDPAT-MLS, BCAL, and LAStools presented the best results, obtaining a low percentage of total errors, 12%, 13%, and 14%, respectively, and a good agreement. However, several tree measurements in the forest stands were classified as ground (Type II error) by many filters such as FUSION (57%) and ALDPAT-MLS (50%).

Finally, LAStools algorithm was the most suitable option in areas dominated by crops and grassland as no errors were committed.

Fig. 7. Effect of point density on filter performance and Cohen's Kappa index (k) (p-value ≤ 0.05).Fig. 8. Effect of scan angle on filter performance and Cohen's Kappa index (k) (p-value ≤ 0.05).

3) *Point Density*: The effect of the variation in point density on the performance of the filter algorithms was assessed quantitatively (Fig. 7). Theoretically, if the resolution of the LiDAR data decreases, it is harder to separate the bare earth from the objects [29]. This fact results in more susceptibility to commit Type II errors on low-resolution DEMs as they are surrounded closely by ground points [30]. However, our results do not allow us to draw firm conclusions on the effect of the point density on the filter performance as also Sithole and Vosselman [29] indicated. In general, it can be seen that the Type II errors increased as did the point density, especially in the case of ALDPAT-PM (6%–62%), although BCAL, FUSION, and ALDPAT-MLS filters did not show a recognizable tendency. On the contrary, Type I errors decreased with increasing point density in all filters tested. LAStools had the lowest total error (12% with a very good Kappa index of agreement, 0.74) in areas with point

densities of 1 point/m². In short, the filters show different responses to the variations in point densities.

4) *Scan Angle*: Experience in the past revealed that artifacts attributed to multipathing, i.e., returns located well below the ground, often occur when scan angles exceed 12° – 14° , particularly in dense forest stands [61], [62]. PNOA mission is reluctant to discard those returns as the resulting return density nears the minimum specified. Therefore, the effect of scan angle on filter performance was examined. As can be seen in Fig. 8, error rates in filtering tend to increase when the scan angle is higher than $\pm 14^\circ$. However, there are exceptions. ALDPAT-PM reduced their Total and Type II errors, but increased the Type I errors (5%–23%). This is similar in ALDPAT-ETW filter, which dramatically increased Type I error from 16% to 46%. In contrast, the rest of methods followed a different trend, increasing their error with scan angles higher than $\pm 14^\circ$. Therefore,

it can be assumed that the ALDPAT-PM and ALDPAT-ETEW filters are less sensitive to errors introduced by the increment in scan angle, despite of their inconspicuous Kappa index. On the opposite, interpolation-based filters resulted in increased misclassifications, particularly Type I errors.

IV. DISCUSSION

LiDAR technology is recognized as a cost-effective mean of sampling the terrain surface, particularly over extensive areas, and it is increasingly becoming very important in the estimation of forest structural parameters [6].

Spain has joined the short list of countries, such as Netherlands, Belgium, Switzerland, or some parts of the United States, that have conducted national-level LiDAR programs with the purpose of providing national high-resolution DEMs (2–5 m), using this active remotely sensed data [14]. These national DEMs are typically produced by the same national agencies that collect the LiDAR data, and in those, the cost for commercial software to filter the point clouds is only marginal compared to the cost of data collection. However, this can be an issue for smaller organizations or research institutes that collect LiDAR data. This raises the question. For users who do not have access to a commercial product to filter LiDAR data, what other tools for point cloud classification are accessible?

Filtering of LiDAR point clouds into terrain and off-terrain datasets is critical in developing accurate surface models [31], [59], but the main drawback is the large number of methods developed to process the data. Several comparative reviews of some of these existing filters report strengths and weaknesses of each one [27], [29]. However, no comprehensive comparative review of nonproprietary tools (e.g., [44] and [45]) and available filtering software has been made to guide potential users of PNOA-LiDAR data in selecting a proper method.

This study tried to assess the potential of seven filtering algorithms in the classification of discrete return PNOA-LiDAR point clouds, in a Mediterranean pine forest environment. Our goal was to assist users in the selection of the best-processing method for a typical Mediterranean landscape with a variety of covers and terrain characteristics.

Ranges for the initial parameters included in the different approaches were selected by reviewing the literature of the algorithm developers, considering the study-area characteristics and comparing unfiltered and filtered results iteratively. In this sense, defining an appropriate threshold is even more important when dealing with slope-based methods [36].

In this research, the reference data used to check the filtering quality were generated by hand, classifying every point into ground and nonground ones, with the support of LiDAR-derived products, orthophotography, and fieldwork information. Optimum filters and parameters were selected through a qualitative (3-D shaded relief visualizations) and quantitative (error metrics and Kappa index) analysis of the seven algorithms. Furthermore, error associated with LiDAR classification on different landscapes and data contexts, i.e., terrain slope, land cover type, point density, and scan angle of LiDAR data, was assessed.

Our results agree with the conclusions of Sithole and Vosselman [29], who pointed out that filters are not foolproof, and the best filter algorithm and its parameters may vary from one scene condition to another. The absence of severe biases in classified ground returns (success rate of 83.3% with a Kappa index of 0.67) suggests that MCC appropriately balanced Type II and Type I errors [19]. The MCC algorithm, designed for forest landscapes, performs exceptionally well with regard to Type I errors (see Table V), thus maximizing the number of classified ground returns and increasing the detail in the bare-earth surface. For instance, footpaths and roads [Fig. 4(b)] tend to be retained very well [19]. However, in future, the research should address Type II errors, as it mistakenly classifies objects as ground, thus creating bulges and other artifacts [37]. In addition, processing time is a current drawback in the implementation of the MCC method [19].

The MCC and LAsTools algorithms were successful, suggesting that novice users can achieve good results with them using minimal parameters. As Hu *et al.* [59] and Mongus and Žalik [36] noted, the sensitivity to parameters is the major downside for ALDPAT morphological-based filters, which adds value to the interpolated-based methods, such as MCC or LAsTools. In fact, these two filters were the least sensitive to the presence of points on slopes higher than 15°, as indicated their total errors around 20% and the good agreement obtained (Kappa about 0.60).

ALDPAT-MLS filter, a method based on slope or height differences between neighboring points [20], was unreliable due to the large Type II errors obtained, so further investigation about the selection of a threshold to separate ground and nonground points will be needed [30]. In this method, there is still a great uncertainty about whether the differences in elevation of points are caused by the presence of objects or by variations in the terrain altitudes, especially on abrupt surfaces [28].

The results of the ALDPAT morphology-based filters showed in Table V agree with those obtained by Zhang and Whitman [30], who concluded that the PM filter produced the least error among the three methods. A common weakness of morphological methods, apart from the significant loss of information associated with the conversion of the point cloud into a raster image to perform morphological operations [36], is the difficulty in maintenance of terrain features when the window sizes are changed with the operations of openings and closings [25], [26], [28], [31]. The selection of window sizes is critical to remove suitably objects with different sizes [27], which implies using additional knowledge about the extent of objects in the study area [25], [30]. This fact prevents morphological methods from being fully automated [29], [40], [63].

In accordance with Kraus and Pfeifer [5], poor results were achieved with the linear-prediction-based method implemented in FUSION, since it uses a rough surface approximation for determining a buffer zone within which points are classified as bare earth, not preserving terrain details and misclassifying small objects [40]. The major difficulty in this method relies on the selection of customizable parameters to define the intermediate surface and a suitable threshold to classify points [36].

Vegetation density largely determines the ratio of ground to vegetation returns [19]. MCC filter performed the best with low vegetation conditions, whereas the ALDPAT-MLS and BCAL methods were somewhat more suitable for filtering the Aleppo pine forest in terms of total error, although the ALDPAT-MLS filter showed a 50% of Type II error. As Tinkham *et al.* [44] noted in their comparative study between MCC, BCAL, and another custom filter, BCAL is able to create a more reliable surface in very dense, continuous vegetation like those areas of the *Pinus halepensis* Mill. forest included in the study sites. This is most likely due to the block minimum approach that BCAL uses, allowing the creation of a surface from fewer points than MCC. On the other hand, LAsTools was the only filter not committing misclassification in areas covered by crops and grasslands. In this way, users might consider the possibility of combining multiple classification procedures to exploit the strengths of each, depending on the cover type. Thus, our research supports the idea of Sithole and Vosselman [29], who suggested testing the environment to be filtered to avoid unpredictable results.

Empirical analysis regarding the selection of filter parameters in forest studies is necessary to determine which are optimal [44], but also more comparative studies focusing on open-source point classification algorithms, which are more accessible. In addition, future research should conduct a more thorough accuracy analysis in areas of higher complexity, as those occupied by low vegetation, to provide a clearer guidance to specific users [27]. In this sense, the new generation of airborne laser scanning sensors, the so-called full-waveform scanners, offer further information about the targets included in the footprint than location alone: peak amplitude which relates to radiometric properties of the target, and pulse width, which is a measure of surface roughness and slope. This additional information might be very useful to discriminate low vegetation from bare-earth reflections and might help to diminish the problems of discrete airborne LiDAR systems prior to DEM generation [64].

Additionally, filter developers are encouraged to provide more detailed information about filtering processes and parameterizations to facilitate the LiDAR data management [59].

As Sithole and Vosselman [29] pointed out, the effects of the reduced point density are most likely minor compared to the errors introduced by the complexity of the scene, such as low vegetation on steep slopes. Zhang and Whitman [30] concluded that point density has relatively less effect on the filtering results. Our results demonstrate that most of the filters can well identify bare-earth points despite the low count of bare-earth points. The complexity of the test sites determined that even at the highest resolutions, i.e., higher than 2 points/m², the filters encounter difficulties, which then masks the performance of the filters at lower resolutions [29]. Thus, the results showed that point densities up to 1 point/m² did not necessarily cause lower accuracies in the filtering process. It should be noted that the effect of point density on filter performance is also influenced by vegetation types (crops and grassland, scrubs, or forest). The point density is not only determined by the amount of emitted laser pulses but also by the presence or absence of vegetation. In forest areas, the interaction of laser pulses with

tree crowns, branches, and leaves is higher, increasing the point density available for filtering and the complexity of the data.

In the case of the scanning angle, it has been shown that the error in the filtering process increases with scan angle in the progressive TIN densification and in the interpolation-based filters, which had good Kappa index of agreement. However, the morphological filters do not show this tendency.

In order to complement the results presented in this study, it would be useful to focus future research on the analysis of the error distribution in DEM generation, as this error can be propagated subsequently into derived products [16], [30]. As shown in Figs. 3 and 4, a ridge or a hilltop that is locally higher than other portions of the ground surface may resemble an above-ground object and be classified as such [31], or terrain points in steep slopes may be classified as vegetation as lie at the same height, increasing Type II errors. This poses a challenge to the surface-based filters used in this study [31]. The preservation of discontinuous ground features such as pine terrace plantations, frequent in the Mediterranean landscapes, should be better examined. In this sense, Mongus and Žalik [36] argue that preserving ridges may sometimes be unsuccessful when the surface model is created, because the interpolation procedures tend to estimate smooth terrain.

In conclusion, although most of the technical hardware difficulties have been solved, the development of more accurate algorithms and methods for modeling of LiDAR data is still necessary [31].

V. CONCLUSION

This paper proposes a comparison of seven filtering methods, implemented in nonproprietary tools and openly available software, for bare-earth extraction from PNOA-LiDAR data in a Mediterranean forest landscape. These methods, designed for filtering discrete return data, include the progressive TIN densification (LAsTools), the weighted linear least-squares interpolation-based (FUSION), the multiscale curvature classification (MCC), the interpolation-based (BCAL), the elevation threshold with expand window (ALDPAT-ETEW), the progressive morphological (ALDPAT-PM), and the maximum local slope (ALDPAT-MLS) filters. As confirmed our results, a high level of accuracy is achieved by the MCC algorithm (83.3% of success rate and Kappa index of 0.67), as well as by the LAsTools (82.8% of success rate and Kappa index of 0.66), compared with the rest of filtering algorithms tested. However, each method has its strengths and weaknesses as none of them worked perfectly, and all methods were susceptible to both Type I and Type II errors, due to the complexity of the earth surface, i.e., terrain slopes and cover types, and the point density and scan angle, to a lesser extent. MCC and LAsTools filters produced good results, but they differed in Type I/Type II error counting. Sprouted scrub, stumps, and woody debris were the more problematic cover type in filtering, as well as terrain slopes higher than 15°. However, less firm conclusions can be drawn from point density and scan angle variables, because morphological methods are less sensitive to these factors. The quality of the LiDAR-derived DEMs directly affects the quality of the LiDAR-derived canopy heights used in subsequent

vegetation modeling, where the removal of nonground points from the raw data is a critical stage. Filtering parameters implemented in this study and the results obtained establish the first baseline for potential users of medium–low density point clouds as the PNOA-LiDAR mission, in the absence, until now, of information describing in detail the suitability of parameters and filtering methods in areas occupied by Aleppo pine forest mixed with evergreen shrub.

ACKNOWLEDGMENT

The authors are grateful to the Training Center (CENAD) “San Gregorio” for assistance in the field. The LiDAR data were provided by the National Center for Geographic Information of Spain (CNIG). The authors also would like to thank two anonymous reviewers, each of whom provided thorough and thoughtful comments that directed important improvements in the manuscript.

REFERENCES

- [1] E. J. Huising and L. M. Gomes Pereira, “Errors and accuracy estimates of laser data acquired by various laser scanning systems for topographic applications,” *ISPRS J. Photogramm. Remote Sens.*, vol. 53, no. 5, pp. 245–261, 1998.
- [2] M. S. Renslow, *Manual of Airborne Topographic LiDAR*. Bethesda, MD, USA: ASPRS, 2012.
- [3] E. P. Baltsavias, “A comparison between photogrammetry and laser scanning,” *ISPRS J. Photogramm. Remote Sens.*, vol. 54, nos. 2–3, pp. 83–94, 1999.
- [4] K. A. Razak, M. W. Straatmaa, C. J. van Westen, J.-P. Malet, and S. M. de Jong, “Airborne laser scanning of forested landslides characterization: Terrain model quality and visualization,” *Geomorphology*, vol. 126, pp. 186–200, 2011.
- [5] K. Kraus and N. Pfeifer, “Determination of terrain models in wooded areas with airborne laser scanner data,” *ISPRS J. Photogramm. Remote Sens.*, vol. 53, no. 4, pp. 193–203, 1998.
- [6] G. Vosselman and H. G. Maas, *Airborne and Terrestrial Laser Scanning*. Dunbeath, Scotland: Whittles, 2010.
- [7] C. Chen, Y. Li, W. Li, and H. Dai, “A multiresolution hierarchical classification algorithm for filtering airborne LiDAR data,” *ISPRS J. Photogramm. Remote Sens.*, vol. 83, pp. 1–9, 2013.
- [8] E. Ahokas, H. Kaartinen, and J. Hyyppä, “A quality assessment of airborne laser scanner data,” *Int. Arch. Photogramm. Remote Sens. Spat. Inf. Sci.*, vol. 34, Part 3, p. W13, 2003.
- [9] C. Hopkinson, “The influence of flying altitude, beam divergence, and pulse repetition frequency on laser pulse return intensity and canopy frequency distribution,” *Can. J. Remote Sens.*, vol. 33, no. 4, pp. 312–324, 2007.
- [10] M. E. Hodgson *et al.*, “An evaluation of LiDAR-derived elevation and terrain slope in leaf-off conditions,” *Photogramm. Eng. Remote Sens.*, vol. 71, no. 7, pp. 817–823, 2005.
- [11] J. Su and E. Bork, “Influence of vegetation, slope, and LiDAR sampling angle on DEM accuracy,” *Photogramm. Eng. Remote Sens.*, vol. 72, no. 11, pp. 1265–1274, 2006.
- [12] C. W. Bater and N. C. Coops, “Evaluating error associated with LiDAR-derived DEM interpolation,” *Comput. Geosci.*, vol. 35, no. 2, pp. 289–300, 2009.
- [13] J. P. Wilson, “Digital terrain modeling,” *Geomorphology*, vol. 137, pp. 107–121, 2012.
- [14] C. Vêga, S. Durrieu, J. Morel, and T. Allouis, “A sequential iterative dual-filter for LiDAR terrain modeling optimized for complex forested environments,” *Comput. Geosci.*, vol. 44, pp. 31–41, 2012.
- [15] P. F. Fisher and N. J. Tate, “Causes and consequences of error in digital elevation models,” *Prog. Phys. Geogr.*, vol. 30, no. 4, pp. 467–489, 2006.
- [16] C. Hutton and R. Brazier, “Quantifying riparian zone structure from airborne LiDAR: Vegetation filtering, anisotropic interpolation, and uncertainty propagation,” *J. Hydrol.*, vols. 442–443, pp. 36–45, 2012.
- [17] G. Sohn and I. Dowman, “Terrain surface reconstruction by the use of tetrahedron model with the MDL criterion,” *Int. Arch. Photogramm. Remote Sens. Spat. Inf. Sci.*, vol. 34, Part 3A, pp. 336–344, 2002.
- [18] A. Kobler, N. Pfeifer, P. Ogrinc, L. Todorovski, K. Ostir, and S. Dzeroski, “Repetitive interpolation: A robust algorithm for DTM generation from aerial laser scanner data in forested terrain,” *Remote Sens. Environ.*, vol. 108, no. 1, pp. 9–23, 2007.
- [19] J. S. Evans and A. T. Hudak, “A multiscale curvature algorithm for classifying discrete return LiDAR in forested environments,” *IEEE Trans. Geosci. Remote Sens.*, vol. 45, no. 4, pp. 1029–1038, Apr. 2007.
- [20] G. Vosselman, “Slope based filtering of laser altimetry data,” *Int. Arch. Photogramm. Remote Sens. Spat. Inf. Sci.*, vol. 33, Part B3, pp. 935–942, 2000.
- [21] S. Filin and N. Pfeifer, “Segmentation of airborne laser scanning data using a slope adaptive neighbourhood,” *ISPRS J. Photogramm. Remote Sens.*, vol. 60, no. 2, pp. 71–80, 2006.
- [22] C. K. Wang and Y. H. Tseng, “DEM generation from airborne LiDAR data by an adaptive dual-directional slope filter,” *Int. Arch. Photogramm. Remote Sens. Spat. Inf. Sci.*, vol. 38, Part 7B, pp. 628–632, 2010.
- [23] D. Tóvári and N. Pfeifer, “Segmentation based robust interpolation—A new approach to laser data filtering,” in *Proc. ISPRS Workshop Laser Scann. (WG III/3, III/4, V/3)*, Enschede, The Netherlands, 2005, 6 pp.
- [24] J. Kilian, N. Haala, and M. Englich, “Capture and evaluation of airborne laser scanner data,” *Int. Arch. Photogramm. Remote Sens.*, vol. 31, Part B3, pp. 383–388, 1996.
- [25] K. Zhang, S. C. Chen, D. Whitman, M. L. Shyu, J. Yan, and C. Zhang, “A progressive morphological filter for removing non ground measurements from airborne LiDAR data,” *IEEE Trans. Geosci. Remote Sens.*, vol. 41, no. 4, pp. 872–882, Apr. 2003.
- [26] Q. Chen, P. Gong, D. Baldocchi, and G. Xie, “Filtering airborne laser scanning data with morphological methods,” *Photogramm. Eng. Remote Sens.*, vol. 73, no. 2, pp. 175–185, 2007.
- [27] X. Meng, N. Currit, and K. Zhao, “Ground filtering algorithms for airborne LiDAR data: A review of critical issues,” *Remote Sens.*, vol. 2, no. 3, pp. 833–860, 2010.
- [28] Y. Li, H. Wu, H. Xu, R. An, J. Xu, and Q. He, “A gradient-constrained morphological filtering algorithm for airborne LiDAR,” *Opt. Laser Technol.*, vol. 54, pp. 288–296, 2013.
- [29] G. Sithole and G. Vosselman, “Experimental comparison of filter algorithms for bare-earth extraction from airborne laser scanning point clouds,” *ISPRS J. Photogramm. Remote Sens.*, vol. 59, nos. 1–2, pp. 85–101, 2004.
- [30] K. Zhang and D. Whitman, “Comparison of three algorithms for filtering airborne LiDAR data,” *Photogramm. Eng. Remote Sens.*, vol. 71, no. 3, pp. 313–324, 2005.
- [31] J. Zhang and X. Lin, “Filtering airborne LiDAR data by embedding smoothness-constrained segmentation in progressive TIN densification,” *ISPRS J. Photogramm. Remote Sens.*, vol. 81, pp. 44–59, 2013.
- [32] P. Axelsson, “DEM generation from laser scanner data using adaptive TIN models,” *Int. Arch. Photogramm. Remote Sens.*, vol. 33, Part B4/1, pp. 110–117, 2000.
- [33] J. L. Silván-Cárdenas and L. Wang, “A multi-resolution approach for filtering LiDAR altimetry data,” *ISPRS J. Photogramm. Remote Sens.*, vol. 61, no. 1, pp. 11–22, 2006.
- [34] G. Y. Lu and D. W. Wong, “An adaptive inverse-distance weighting spatial interpolation technique,” *Comput. Geosci.*, vol. 34, no. 9, pp. 1044–1055, 2008.
- [35] X. Meng, L. Wang, J. L. Silván-Cárdenas, and N. Currit, “A multi-directional ground filtering algorithm for airborne LiDAR,” *ISPRS J. Photogramm. Remote Sens.*, vol. 64, no. 1, pp. 117–124, 2009.
- [36] D. Mongus and B. Žalik, “Parameter-free ground filtering of LiDAR data for automatic DTM generation,” *ISPRS J. Photogramm. Remote Sens.*, vol. 67, pp. 1–12, 2012.
- [37] T. J. Pingel, K. C. Clarke, and W. A. McBride, “An improved simple morphological filter for the terrain classification of airborne LiDAR data,” *ISPRS J. Photogramm. Remote Sens.*, vol. 77, pp. 21–30, 2013.
- [38] C. Brenner, “Towards fully automatic generation of city models. International archives of the photogrammetry,” *Remote Sens. Spat. Inf. Sci.*, vol. 33, Part 3B, pp. 85–92, 2000.
- [39] W. Zhou and A. Troy, “An object-oriented approach for analysing and characterizing urban landscape at the parcel level,” *Int. J. Remote Sens.*, vol. 29, Part 11, pp. 3119–3135, 2008.
- [40] M. Bartels and H. Wei, “Threshold-free object and ground point separation in LiDAR data,” *Pattern Recognit. Lett.*, vol. 31, no. 10, pp. 1089–1099, 2010.
- [41] M. Bartels and H. Wei, “Segmentation of LiDAR data using measures of distribution,” *Int. Arch. Photogramm. Remote Sens. Spat. Inf. Sci.*, vol. 36, Part 7, pp. 426–431, 2006.

- [42] Y. Bao *et al.*, "Classification of LiDAR point cloud and generation of DTM from LiDAR height and intensity data in forested area," *Int. Arch. Photogramm. Remote Sens. Spat. Inf. Sci.*, vol. 37, Part 3B, pp. 313–318, 2008.
- [43] A. B. Jahromi, M. J. V. Zoej, A. Mohammadzadeh, and S. Sadeghian, "A novel filtering algorithm for bare-earth extraction from airborne laser scanning data using an artificial neural network," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 4, no. 4, pp. 836–843, Dec. 2011.
- [44] W. T. Tinkham *et al.*, "A comparison of two open source LiDAR surface classification algorithms," *Remote Sens.*, vol. 3, pp. 638–649, 2011.
- [45] K. Korzeniowska, N. Pfeifer, G. Mandlbürger, and A. Lugmayr, "Experimental evaluation of ALS point cloud ground extraction tools over different terrain slope and land-cover types," *Int. J. Remote Sens.*, vol. 35, pp. 4673–4697, 2014.
- [46] S. M. Vicente-Serrano, T. Lasanta, and C. Gracia, "Aridification determines changes in forest growth in Pinus halepensis forests under semiarid Mediterranean climate conditions," *Agric. For. Meteorol.*, vol. 150, no. 4, pp. 614–628, 2010.
- [47] E. Notivol, A. Cabanillas, R. González, and C. Revuelta, "Caracterización de masas naturales de pino carrasco (*Pinus halepensis* Mill.) en la Depresión del Ebro," in *Proc. 4º Congreso Forestal Español La ciencia Forestal: Respuestas para la Sostenibilidad*, 2005, 7 pp.
- [48] M. Tanase, J. de la Riva, and F. Pérez-Cabello, "Estimating burn severity at the regional level using optically based indices," *Can. J. For. Res.*, vol. 41, no. 4, pp. 863–872, 2011.
- [49] M. Tanase, J. de la Riva, M. Santoro, F. Pérez-Cabello, and E. Kasischke, "Sensitivity of SAR data to post-fire forest regrowth in mediterranean and boreal forests," *Remote Sens. Environ.*, vol. 115, pp. 2075–2085, 2011.
- [50] E. Chuvieco, *Earth Observation of Wildland Fires in Mediterranean Ecosystems*. Berlin, Germany: Springer-Verlag, 2009, 257 pp.
- [51] P. Axelsson, "Processing of laser scanner data—Algorithms and applications," *ISPRS J. Photogramm. Remote Sens.*, vol. 54, nos. 2–3, pp. 138–147, 1999.
- [52] R. McGaughey, *FUSION/LDV: Software for LIDAR Data Analysis and Visualization*. Seattle, WA, USA: US Dept. Agriculture, Forest Service, Pacific Northwest Research Station, 2009, 175 pp.
- [53] K. Kraus and E. Mikhail, "Linear least squares interpolation," *Photogramm. Eng.*, vol. 38, pp. 1016–1029, 1972.
- [54] A. S. Maguaya, V. Junttila, and T. Kauranne, "Adaptive algorithm for large scale DTM interpolation from LiDAR data for forestry applications in steep forested terrain," *ISPRS J. Photogramm. Remote Sens.*, vol. 85, pp. 74–83, 2013.
- [55] D. Streutker and N. Glenn, "LiDAR measurement of sagebrush steppe vegetation heights," *Remote Sens. Environ.*, vol. 102, nos. 1–2, pp. 135–145, 2006.
- [56] K. Zhang and Z. Cui, *Airborne LIDAR Data Processing and Analysis Tools. ALDPAT 1.0*. Miami, FL, USA: National Center for Airborne Laser Mapping, 2007, 81 pp.
- [57] P. Soille, *Morphological image analysis: Principles and applications*. New York, NY, USA: Springer, 2003, 392 pp.
- [58] J. Cohen, "A coefficient of agreement for nominal scales," *Educ. Psychol. Meas.*, vol. 20, pp. 37–46, 1960.
- [59] H. Hu *et al.*, "An adaptive surface filter for airborne laser scanning point clouds by means of regularization and bending energy," *ISPRS J. Photogramm. Remote Sens.*, vol. 92, pp. 98–111, 2014.
- [60] J. R. Landis and G. G. Koch, "The measurement of observer agreement for categorical data," *Biometrics*, vol. 33, pp. 159–174, 1977.
- [61] D. Gatzliolis and H. E. Andersen, "A guide to LiDAR data acquisition and processing for the forests of the Pacific Northwest," US Dept. Agriculture, Forest Service, Pacific Northwest Research Station, Portland, OR, USA, General Tech. Rep. PNW-GTR-768, 2008, 32 pp.
- [62] J. Holmgren and M. Nilsson, "Simulating the effects of LiDAR scanning angle for estimation of mean tree height and canopy closure," *Can. J. Remote Sens.*, vol. 29, no. 5, pp. 623–632, 2003.
- [63] X. Liu, "Airborne LiDAR for DEM generation: Some critical issues," *Prog. Phys. Geogr.*, vol. 32, no. 1, pp. 31–49, 2008.
- [64] G. Vosselman and H. G. Maas, *Airborne and Terrestrial Laser Scanning*. Dunbeath, Scotland: Whittles, 2010.



Antonio Luis Montealegre was born in Barcelona, Spain, in 1986. He received the Bachelor's degree in geography and the Master's degree in geographic information technologies for land management: geographic information systems (GIS) and remote sensing from the University of Zaragoza, Zaragoza, Spain, in 2009 and 2010, respectively. He is currently pursuing the Ph.D. degree in geography and land management at the same university.

Since March 2011, his work has been financed by the Government of Aragón, Department of Science, Technology and University, being involved in the activities of the GEOFOREST Research Group, Environmental Sciences Institute, University of Zaragoza. His research interests include LiDAR remote sensing data processing techniques, applied particularly to fire severity assessment, forest inventory, fuel type modeling, and DEMs generation.



María Teresa Lamelas was born in Zaragoza, Spain, in 1977. She received the Ph.D. degrees in geography and in natural sciences from the University of Zaragoza, Zaragoza, Spain, and Damstadt University of Technology, Damstadt, Germany, respectively, in 2007.

From 2008 to 2012, she was an Associated Teacher with the Department of Geography and Land Management, University of Zaragoza. Since 2012, she has been teaching the subject "geographical information systems and remote sensing" and has been a Researcher in the same topic with the Centro Universitario de la Defensa in Zaragoza, Zaragoza, Spain. She is Member of the GEOFOREST Research Group, Environmental Sciences Institute, University of Zaragoza. Her research interests include the use of geographical information technologies in natural resources and hazards modeling.



Juan de la Riva received the Ph.D. degree in geography from the University of Zaragoza, Zaragoza, Spain, in 1994.

He is currently a Professor of Regional Geographic Analysis with the University of Zaragoza, where he coordinates the GIS and remote sensing master course. He is the Head of the GEOFOREST Research Group, Environmental Sciences Institute, University of Zaragoza, and has worked in different management studies in mountain areas, as well as in several research projects, namely, FIRERISK, EROFUEGO, LIGNOSTRUM, RS-FIRE, PIR-FIRE, and FIREGLOBE. His research interests include the study of forested areas, application of GIS and remote sensing techniques, use of remote sensing and GIS for forest characterization, forest fires (risk modeling, fire severity, and postfire environmental dynamic) assessment, and biomass estimation.

3.2 Comparación de métodos de interpolación para la generación de MDE en espacios forestales

RESUMEN

La exploración con un sistema láser aerotransportado permite estimar una gran variedad de parámetros forestales empleando diferentes métricas extraídas de las alturas normalizadas de la nube de puntos a través de un MDE.

En este artículo se han analizado seis rutinas de interpolación considerando diferentes coberturas de suelo y pendientes topográficas para generar una colección de MDE con resoluciones espaciales de 1 y 2 metros. La precisión de los MDE se ha evaluado de dos maneras: usando una muestra extraída de la propia nube de puntos LiDAR, y utilizando un conjunto de 55 puntos de control tomados sobre el terreno con un sistema de posicionamiento global (GPS) de alta precisión.

Se han examinado los efectos de la pendiente del terreno, la cobertura del suelo, la densidad de puntos del suelo y la penetración de los pulsos en el error de interpolación, estratificando el área de estudio con estas variables. Además, se ha utilizado un Árbol de Clasificación y Regresión (CART) que ha permitido el desarrollo de un mapa de incertidumbre en la predicción del MDE para identificar las áreas de peor calidad del modelo.

El método de interpolación “TIN a ráster” ha proporcionado el mejor resultado en el proceso de validación con el conjunto de datos LiDAR, mientras que la rutina *Inverse Distance Weighted* (IDW) ha sido la más precisa tras realizar la validación con los puntos GPS (RMSE de 2,68 cm y RMSE de 37,10 cm, respectivamente).

Article

Interpolation Routines Assessment in ALS-Derived Digital Elevation Models for Forestry Applications

Antonio Luis Montealegre ^{1,*}, María Teresa Lamelas ^{1,2,†} and Juan de la Riva ^{1,†}

¹ GEOFOREST Group, Instituto de Investigación en Ciencias Ambientales de Aragón (IUCA), Department of Geography, University of Zaragoza, Zaragoza 50009, Spain; E-Mails: tlamelas@unizar.es (M.T.L.); delariva@unizar.es (J.R.)

² Centro Universitario de la Defensa de Zaragoza, Zaragoza 50090, Spain

[†] These authors contributed equally to this work.

* Author to whom correspondence should be addressed; E-Mail: monteale@unizar.es; Tel.: +34-876-554-058; Fax: +34-976-761-506.

Academic Editors: Nicolas Baghdadi and Prasad Thenkabail

Received: 30 April 2015 / Accepted: 1 July 2015 / Published: 9 July 2015

Abstract: Airborne Laser Scanning (ALS) is capable of estimating a variety of forest parameters using different metrics extracted from the normalized heights of the point cloud using a Digital Elevation Model (DEM). In this study, six interpolation routines were tested over a range of land cover and terrain roughness in order to generate a collection of DEMs with spatial resolution of 1 and 2 m. The accuracy of the DEMs was assessed twice, first using a test sample extracted from the ALS point cloud, second using a set of 55 ground control points collected with a high precision Global Positioning System (GPS). The effects of terrain slope, land cover, ground point density and pulse penetration on the interpolation error were examined stratifying the study area with these variables. In addition, a Classification and Regression Tree (CART) analysis allowed the development of a prediction uncertainty map to identify in which areas DEMs and Airborne Light Detection and Ranging (LiDAR) derived products may be of low quality. The Triangulated Irregular Network (TIN) to raster interpolation method produced the best result in the validation process with the training data set while the Inverse Distance Weighted (IDW) routine was the best in the validation with GPS (*RMSE* of 2.68 cm and *RMSE* of 37.10 cm, respectively).

Keywords: LiDAR; ALS; interpolation; DEM validation; CART analysis; Mediterranean forest

1. Introduction

Airborne Light Detection and Ranging (LiDAR), also referred to as Airborne Laser Scanning (ALS), is an active remote sensing technology which has gradually become a common tool for collecting elevation information of surface targets with high precision [1,2]. Compared with the traditional photogrammetric method, the accuracies of the small footprint discrete return ALS measurements are unaffected by external light conditions [3]. In addition, the high spatial resolution of ALS outperforms the use of Synthetic Aperture Radar (SAR) [4]. ALS systems are capable of penetrating through vegetation and recording the terrain beneath it [5]. Accordingly, ALS has been widely used for generating accurate and high spatial resolution Digital Elevation Models (DEMs) [6,7] over wide areas [8–12], which is essential for environmental applications.

ALS has already been adopted and accepted as a very valuable tool in forestry due to the three-dimensional nature of data [13]. A wide array of vegetation structural metrics, such as tree height, biomass, crown size, leaf area index, stem volume, basal area, stand density and vertical canopy structure has been estimated (e.g., [14–21]). In this context, estimates are typically based on the height of the point cloud above a continuous gridded DEM representing the bare-Earth's surface [22]. However, the raw ALS data contains a large number of points returned not only from the bare-earth surface but also from other surface objects. These non-ground/object points should be separated or classified, the so-called ALS data filtering, prior to DEM development [13]. This process is the most critical step in DEM generation [23,24], which may also affect the accuracy assessment of the DEMs as some non-ground ALS points may be erroneously labelled as ground points [25].

In addition, the numerous interpolation methods developed to derive a DEM from point data vary widely in their complexity, ease of use, and computational expense, thus presenting their own advantages and disadvantages depending on the characteristics of the data sets [2,26]. The fidelity with which DEMs represent the real surface has been extensively explored in the last decades [27]. However, as Bilskie and Hagen [28] indicated, there is an insufficiency in the literature, as well as in available Geographical Information Systems (GIS) software to efficiently assess the vertical errors related with an interpolation method using ALS data. The selection of an appropriate interpolation algorithm and spatial resolution to generate an accurate DEM becomes an important decision, especially in uneven terrain. In fact, gridding error can comprise a very important, and often neglected, source of inaccuracy in vegetation metrics estimation [29]. This may be especially relevant for canopy height model estimation in forested areas with a low ground-return sample for effective DEM surface interpolation [19,30–32]. In this regard, the density of ground points after filtering a point cloud varies depending on the environmental conditions. For example, in heavily wooded and vegetated areas, ground points will be particularly sparse and the DEM will typically present lower accuracy, detail, and reliability [2,22], because the laser beam penetration through the canopy can be limited.

Many local studies have explored and documented in the last decades that source data density, terrain, land cover type, interpolation method, and grid size affect DEM error [6,28,33–39], although few studies have comprehensively studied the effects of all the aforementioned factors together such as Guo *et al.* [26] or Bater and Coops [22]. Furthermore, it is relevant to study the topographic error associated with ALS data sets of low nominal point density per square meter, captured at national scale, as is the case of the new ALS data provided by the Spanish National Plan for Aerial Orthophotography (PNOA) [40]. Spain has made a significant effort similar to Scandinavia and the USA to provide ALS data. In accordance with the quality levels defined by the 3D Elevation Program (3DEP) of the U.S. Geological Survey National Geospatial Program (NGP) [41], the PNOA-LiDAR project meets the Quality Level 3 (QL3), which implies a vertical accuracy of 20.0 cm *RMSE* and a density of 0.5 points/m². This accuracy aligns with the American Society for Photogrammetry and Remote Sensing (ASPRS) 20-cm Vertical Accuracy Class.

The consideration of error distribution and error propagation in DEMs has been often neglected, perhaps because the immediacy of DEM implementation with tools commonly available in software packages that overrides any concern for accuracy and error [27]. In order to gain a better understanding of the error introduced in DEM development by factors related to the territory and the ALS data acquisition, prediction uncertainty maps can be valuable tools. These maps show the spatial distribution and magnitude of potential error and can assist in the recognition of areas of low quality in the DEMs generated and derived LiDAR products (e.g., [22]).

According to Gatziolis *et al.* [42], nearly all evaluations of the suitability of ALS data for estimating forest structural variables have been carried out in relatively simple forest conditions with a uniform canopy, little if any understory vegetation, and gentle topography [14,15], which probably facilitate the high accuracy of ALS-derived metrics of forest canopy and bare ground extraction. However, such conditions are not common in many forested areas, including the Mediterranean forests of Spain, where little is known about the effects that their complex structure and terrain may have on the suitability of ALS-derived height estimates. Thus, the aim of this research is to assess different interpolation methods in order to generate an optimal DEM to normalize the ALS data captured by the PNOA-LiDAR mission to be used in forestry applications in the context of a study area dominated by Mediterranean pine forests and topographic variability. The research objectives of this paper are to: (1) evaluate the relative performance of six interpolation routines (natural neighbor, Triangulated Irregular Network (TIN) to raster, Inverse Distance Weighted (IDW), Australian National University DEM (ANUDEM), kriging, and point to raster) implemented in ArcGIS 10.1 (ESRI, Redlands, CA, USA) and FUSION 3.30 [43] software; (2) identify the most accurate spatial resolution for DEM creation; (3) assess the effect of terrain slope, land cover, ALS ground point density and pulse penetration on interpolation error; (4) identify the most important variables in error prediction and evaluate the error distribution applying a Classification And Regression Tree (CART) analysis; (5) and provide guidance for users of low density point clouds, as the PNOA-LiDAR, in order to select the more suitable interpolation routine.

2. Methods

2.1. Study Area

The study area consists of two sample sites, T1 (2×2 km) and T2 (4×2 km), located in the central Ebro valley ($41^{\circ}56'N$, $0^{\circ}56'W$), sited northeastern Spain (Figure 1). The Ebro basin constitutes the northernmost semi-arid region in Europe and stretches from the Pyrenees range, in the north, to the Iberian range, in the south.

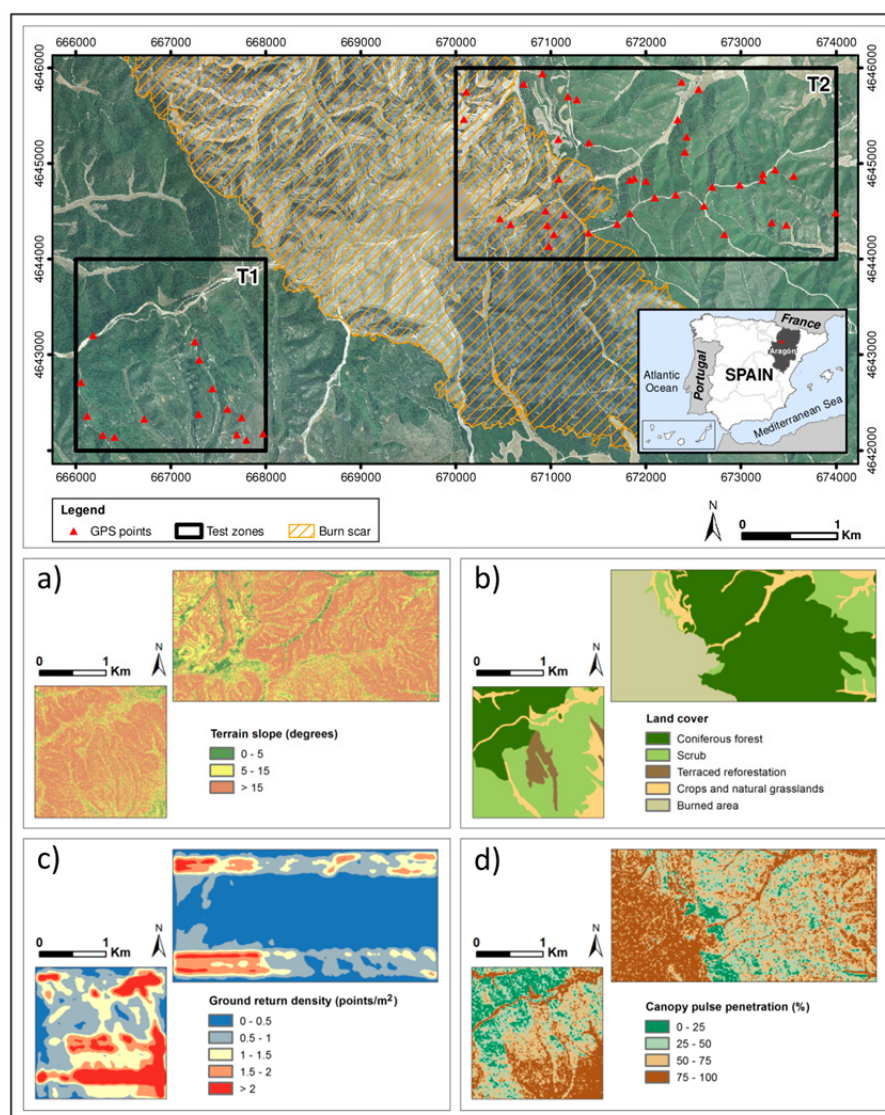


Figure 1. Study area with the two test sites (T1 and T2) and factors influencing DEM accuracy: **(a)** Terrain slope; **(b)** Land cover; **(c)** Ground return density; **(d)** Canopy pulse penetration. The red triangles denote the locations of the reference GPS benchmarks used in the accuracy assessment overlaying a high spatial resolution PNOA-orthophotography captured in 2009.

This area presents a Mediterranean climate with continental features. Annual precipitation averages 350 mm and mostly occurs in autumn and spring. Moreover, the study area presents cold winters, with monthly mean temperature about 7 °C, and hot, dry summers, with temperatures about 24 °C [44]. Topography is characterized by moderate to steep slopes with elevation ranging from 400 m to 750 m.a.s.l.

In the two selected sites, Aleppo pine forests (*Pinus halepensis* Mill.) account for 44% of total cover and pine terrace plantation only 2%. Evergreen shrub vegetation represents 25%, dominated by a mixture of *Quercus coccifera* L., *Juniperus oxycedrus* L. subsp. *macrocarpa* (Sibth. & Sm.) Ball and *Thymus vulgaris* L., and cereal crops cover 10% of the study area. The average height of the canopy is approximately 6.5 m and the average biomass around 45 t/ha. Old stands reach heights of 12–13 m and 90 t/ha [45]. During the last century, the study site has been recurrently affected by fire, with some areas being burned even twice. Two scars caused by wildfires in June 1995 and August 2008, which consumed a total of 5,300 ha of forest, are distinguishable nowadays [46]. Currently, in these areas affected by fire, the vegetation is dominated by shrub species that have colonized rapidly [47]. Thus, this area is characteristic of the distinctive dynamic of the Mediterranean environment, recurrently affected by wildfires [48].

In summary, the selection of T1 and T2 sites was based on the objective of our research, which is to test different interpolation methods to develop DEMs in order to normalize the PNOA-LiDAR data for forestry applications in a typical Mediterranean environment. In this regard, in order to normalize the point heights not only the evaluation of filtering procedures to classify the point cloud is relevant, but also the interpolation methods.

2.2. ALS Data Acquisition

The ALS data were provided by the Spanish National Plan for Aerial Orthophotography (PNOA) [40] and captured in several surveys conducted between 22st January and 5th February 2011, using an airborne Leica ALS60 discrete return sensor. Data were collected with up to four returns measured per pulse, and intensity values from a 1064-nm wavelength laser. The resulting ALS nominal point density was 1.5 point/m² with a vertical accuracy of 0.20 m. Data were delivered in three 2 km × 2 km tiles of raw data points encoded in the ASPRS laser (LAS) binary file format v. 1.1, containing *x*, *y* coordinates (UTM Zone 30 ETRS 1989) and ellipsoidal elevation *z* (ETRS 1989). The properties of the ALS acquisition are summarized in Table 1 [47].

Table 1. Airborne laser scanning (ALS) data specifications and acquisition properties.

Property	Value
Sensor	ALS60
Scanning method	Oscillating plane mirror (saw-tooth pattern)
Date	January and February 2011
Mean flying height above ground (m)	3012
Mean flying speed (km/h)	241
Nominal point density (point/m ²)	1.5
Field of view (degrees)	40
Beam divergence angle (mrad)	0.22

Table 1. Cont.

Property	Value
Scan angle (degrees)	$\pm 22^\circ$
Total extension of LAS files (km ²)	12
Point count	18,495,618
Altimetric accuracy (RMSE in m)	0.20
Planimetric accuracy (RMSE in m)	0.30

2.3. Data Processing

A key step in DEM generation is the previous classification of laser returns as either on or above the ground, filtering out the aboveground returns before interpolating the ground points to generate a surface [7]. In this study, the point cloud classification was performed using the algorithm designed by Evans and Hudak [49] implemented in MCC v.2.1 software. According to Montealegre *et al.* [50], this classification algorithm balances commission (Type II) and omission (Type I) errors and it is very suitable for forested environments. In that study, the relative performance of seven different well known filtering methods not available in proprietary software was evaluated. These methods were the progressive TIN densification algorithm (LASTools), the weighted linear least squares interpolation-based method (FUSION), the multiscale curvature classification (MCC), the interpolation-based filter (BCAL), the elevation threshold with expand window method (ETEW-ALDPAT), the progressive morphological filter (PM-ALDPAT) and the maximum local slope algorithm (MLS-ALDPAT). According to Montealegre *et al.* [50] results, MCC filter presented the lowest overall error (16.7%) and the more problematic cover types in filtering were sprouted scrub, stumps and woody debris, as well as terrain slopes higher than 15° .

2.3.1. Surface Interpolation Methods

Numerous mathematical methods for creating a raster surface from an irregular point cloud exist [13,36]. In this study, we compared several commonly used interpolation methods: Natural neighbor, Triangulated Irregular Network (TIN) to raster, Inverse Distance Weighted (IDW), ANUDEM (Australian National University DEM), kriging, and point to raster. ALS-derived DEMs with spatial resolutions of 1 and 2 m were created for all interpolation routines. A brief overview of these techniques, currently available in GIS software, and their parameterization is presented below and in the Table 2.

The natural neighbor well known as the “area-stealing” or Sibson method finds the closest subset of input samples to an unknown point and applies weights to them based on proportionate areas determined by Voronoi (Thiessen) polygons to interpolate a value [22,26,51,52].

The TIN to raster is based on a set of contiguous, non-overlapping Delaunay triangles to join points in three-dimensional space. Elevation is recorded for each triangle node, while elevations between nodes can be interpolated, thus allowing the generation of a continuous surface [26]. Then, the value of each output raster cell is interpolated from the TIN surface at the center of each cell [2].

The Inverse Distance Weighted (IDW) estimates the cell value by averaging the values of sample data points within its neighborhood [2,22] based on the Tobler’s “first law of geography”. The closer a point is to the center of the cell being estimated, the more influence, or weight, it has in the averaging process [22]. The influence of known points on the interpolated values based on their distance from the

output point can be controlled by defining the power. For example, a power of two is often well suited for deriving raster Canopy Height Models (CHM) considering the shape of the tree canopy and its variation in elevation. However, in order to generate DEMs, different values can be suitable. Bater and Coops [22] obtained better results using a power of three in IDW interpolation to estimate terrain elevation, than using a lower power. The characteristics of the interpolated surface can be controlled by applying a fixed or variable search radius, which limits the number of input points that can be used for calculating each interpolated cell [2,26].

Table 2. Interpolation routines, most important advantages and disadvantages and the parameterizations tested in this study.

Method	Advantages/Disadvantages	Parameterization
Natural neighbor	It is simple because it has no adjustable parameters. It is extremely computationally efficient. It can create artifacts when points are sparse.	Not applicable.
Triangulated Irregular Network (TIN) to raster	It is simple and computationally efficient. If point density is lower than the output cell size, the triangle of the intermediate TIN will be transferred to the output DEM.	Linear and natural neighbor methods were tested to create the raster surface from the TIN.
Inverse Distance Weighted (IDW)	It requires a moderate decision-making and can also be computationally intensive.	Power of 0.5, 1, 1.5, 2, 2.5 and 3, and a variable search radius with 6, 12 and 24 minimum points were tested.
ANUDEM	It allows the incorporation of spatial restrictions in the interpolation process, such as contours, streams, <i>etc.</i> Its primary purpose is to create a surface suitable for hydrologic modeling. It is extremely computationally intensive.	Surfaces were created with drainage enforcement both on and off.
Kriging	It requires a lot of decision-making and it is very computationally intensive.	The fitted model of the semivariogram was “Gaussian”. Sector types of 1, 4, 4 with an offset of 45° and 8, with 2 to 5 neighbors were tested.
Point to raster	It is the simplest method and it is very computationally efficient. Mean is sensitive to extreme values/outliers, especially when the sample size is small.	Not applicable.

The topo to raster or ANUDEM uses an iterative finite difference interpolation technique specifically intended for terrain modelling that more closely represents a natural drainage surface, developed by Hutchinson [53]. Although ANUDEM is capable of incorporating different additional data (e.g., contours or drainage), only ALS ground returns were used for surface development [2,36].

Kriging is an advanced geostatistical procedure that generates an estimated surface from a set of points with z values [2,54,55]. It is based on the regionalized variable theory that assumes that the spatial variation in the phenomenon represented by the z values is statistically homogeneous throughout the surface. In this study ordinary kriging, one of the most commonly applied kriging approaches, was evaluated [2]. It assumes that the variation in elevation values is free of any structural component (drift) [26]. All parameters are determined by weighted least squares methods, which are commonly used to fit semivariogram models [26,56,57].

All the aforementioned methods were implemented in ArcGIS 10.1 software (ESRI, Redlands, CA, USA), while the following was performed with the “Gridsurfacecreate” command included in FUSION LDV 3.30 [43], given the widespread use of this software by researchers in forestry applications. This method identifies the point or points within each output raster cell and assigns an elevation value to the cell based on the averaging z value of those points. If there are no points within the output cell, this is filled using the neighboring cell heights.

2.3.2. DEM Accuracy Assessment

ALS ground returns were randomly divided into prediction (training) and validation (test) data sets, consisting of 7,585,872 (80%) and 1,896,468 (20%) points, respectively. The training data set percentage was selected in order to ensure the generation of DEMs with 1 and 2 m resolutions, while the test data set was used to assess vertical errors in elevation without compromising the integrity of the ALS data [28]. In this sense, the ALS ground returns in the prediction data set presented a nominal point density of 1.3 points/m² in T1 and 0.6 points/m² in T2, corresponding to a point spacing of 0.86 and 1.32 m, respectively. It was not intended to test the absolute geodetic accuracy of the DEMs because the validation data were subject to the same degree of positional error as the prediction data, *i.e.*, less than 0.30 m [22,28].

Additionally, a complementary validation was performed using a finite sample of 55 high-accuracy geodetic control points collected with the Leica VIVA GS15 CS10 GNSS real-time kinematic (RTK) global positioning system and located randomly but ensuring covering the whole variability of the study area (see location of the points in Figure 1). The 49% of the ground control points were taken on pine forest, 22% on scrub, 18% on a burned area and the remaining 10% on crops and grasslands.

2.3.3. Error Analysis

Once the ALS-DEMs were developed using the prediction data set, the validation ALS data set and ground truth GPS checkpoints were used to compare the biases and accuracies of the surfaces [39]. The vertical error of every point in the validation data with respect to the predicted value in each DEM, generated with different methods at 1 and 2 m resolution, was calculated using the following Equation (1):

$$E(x, y) = P_z(x, y) - M_z(x, y) \quad (1)$$

where E is the error at location (x, y) , P_z is the predicted value of the DEM at location (x, y) , and M_z is the measured value from the validation data, both ALS and GPS data set at location (x, y) .

Furthermore, other global statistics to assess the overall performance of the interpolation routines, such as *Mean Error (ME)*, *Mean Absolute Error (MAE)* and *Root Mean Square Error (RMSE)*, were computed [22,28,33,37]. See Equations (2)–(4):

$$ME = \frac{1}{n} \sum_{i=1}^n (P_z(x, y) - M_z(x, y)) \quad (2)$$

$$MAE = \frac{1}{n} \sum_{i=1}^n (|P_z(x, y) - M_z(x, y)|) \quad (3)$$

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^n (P_z(x, y) - M_z(x, y))^2} \quad (4)$$

where n is the number of test points.

In addition, with the aim of assessing the effect on interpolation errors of terrain slope, land cover structure, return density and canopy pulse penetration, following Bater and Coops [22] approach, *MAEs* derived from the validation ALS data set were summarized across the range of values of each variable (Figure 1). It should be noted that this analysis was not performed with 55 GPS points because the sample is small to ensure statistical validity of the results.

In order to examine the effects of terrain slope on the error, a slope steepness model was derived from all ALS ground returns in the data set. This slope map was reclassified into three categories, specifically: 0°–5° gentle; 5°–15° moderate; >15° steep slopes (Figure 1a).

A structural analysis of the variability of land cover existing in T1 and T2 sites was performed using the land cover map from the CORINE (Coordination of Information on the Environment) program of the European Commission actualized to 2012. Five structural classes were obtained in the study area (Figure 1b): Coniferous forest, terraced reforestation, burned area, scrub, and crops and natural grassland.

With respect to the influence of ALS return density, a raster surface was generated computing the number of ground returns within each pixel of 1 m². This continuous variable was then discretized into 5 ranges of ground return density: 0–0.5, 0.5–1, 1–1.5, 1.5–2, >2 points/m² (Figure 1c).

As for canopy pulse penetration, a canopy closure model was derived from the ALS data set using the “Cover” command included in FUSION LDV 3.30. [43]. The proportion of the pulses that penetrate canopy and reach the ground was calculated using a 15 m × 15 m grid and a ground tolerance of 1 meter. Then, canopy pulse penetration was categorized into 4 classes: 0%–25%, 25%–50%, 50%–75% and 75%–100% (Figure 1d).

Finally, in order to analyze to what an extent the differences in the error obtained by each interpolation methods are statistically significant, a Kruskal Wallis analysis was performed.

2.3.4. Error Prediction

A CART analysis [58,59] performed with IBM SPSS Statistics 20 was used as an exploratory data technique to uncover those variables having the most influence on DEM error and to examine the spatial distribution of the prediction uncertainty. This is interesting and necessary information to control the prediction uncertainty in ALS derived products that could be used by forest managers. Following Bater and Coops [22] approach, *MAE* computed previously with the test data set (20%) was related to the variables described in the previous section. CART is a non-parametric modeling approach that can explain the response of a dependent variable from a set of independent continuous and/or categorical variables using binary recursive partitioning of the data [59]. This leads to increasingly homogeneous subsets, based on independent variable splitting criteria using variance minimizing algorithms. The dependent data are partitioned into a series of descending left and right child nodes derived from parent nodes. Once the partitioning has concluded, the child nodes are designated as terminal nodes in which all cases have the same value for the dependent variable, *i.e.*, they are homogeneous or “pure” nodes and do not require further splitting [58]. CART is a procedure of data mining tools widely used in land use

change modeling (e.g., [60]) and commonly applied to remote sensing data as a rule-based classification design (e.g., [61–63]). The output from a CART analysis is a series of logical if-then conditions ending in terminal nodes predicting the value of the response variable [22]. These conditional rules, generated by the decision tree, were implemented in ArcGIS 10.1 software (ESRI, Redlands, CA, USA) using the raster layers presented in Figure 1, in order to produce a categorical map of prediction uncertainty.

3. Results

3.1. Error Analysis

Global statistics for the DEM validation using the ALS test data set and the GPS benchmarks are presented in Tables 3 and 4, respectively. The *ME* shown in Table 3 was sub-centimeter using all interpolation algorithms with both spatial resolutions (1 and 2 m), with the exception of point to raster interpolation, which underestimates the prediction by more than 1 cm. Overall, the results show that interpolation biases were negligible in accordance with Bater and Coops [22] and Gallay *et al.* [64]. Analyzing the results in more detail, with the exception of point to raster model, the DEMs derived with the rest of methodologies underestimate ground elevations at 1 m resolution, while ground elevations were over-predicted at 2 m resolution. *RMSEs* ranged from 2.68 to 17.67 cm, decreasing for all interpolation routines with an increase in spatial resolution from 2 to 1 m. *MAE* was also very consistent with respect to the method used, presenting a higher variability in both spatial resolutions. For instance, the best interpolation routine was TIN to raster with a *MAE* of 11.73 cm in 1 m resolution DEM, which increased to 16.94 cm in 2 m resolution DEM. Therefore, the smaller the grid sizes, the lower *MAE* and *RMSE* in accordance with the results obtained by Bilskie and Hagen [28]. The highest range of error was obtained by kriging, point to raster and natural neighbor interpolation methods, while ANUDEM, IDW and TIN to raster presented the lowest ones, below 7 m.

In general, the results of Table 3 show that TIN to raster interpolation method is the optimal solution in 1 m resolution DEM generation, presenting the lowest *RMSE*, *MAE* and *SD* values (2.68 cm, 11.73 cm and 0.16 m, respectively). However, natural neighbor and ANUDEM obtained also good results with lower *RMSE* and *MAE* values, in comparison to IDW, kriging and point to raster interpolation methods. In fact, kriging method performed slightly better in DEMs generated at 2 m cell size, being 2.30 cm more accurate than TIN to raster method. The main drawback of kriging methodology is the processing time required which is almost three times longer than in IDW interpolation routine, as well as the flexibility of kriging, which can require a lot of decision-making [2]. It should be noted that ANUDEM methodology performed relatively well in both resolutions considering that drainage enforcement was not applied.

Table 4 shows the overall vertical error statistics obtained with the 55 checkpoints captured with the high precision GPS. According to Liu [25] and considering the difficult field conditions, the GPS observations presented a high accuracy (vertical and horizontal accuracy of 2.38 and 1.32 cm, respectively). The difference between survey elevations and the ALS-derived DEM elevations was higher (around 30 cm) in terms of *ME*, compared to the values obtained previously and reported in the Table 3. It should be noted that the GPS sample is not affected by the horizontal and vertical errors that present the ALS training data set (see Table 1). This fact contributes significantly to the error detected

with the reference GPS benchmarks validation. Although the 55 survey points represent a smaller independent sample, this approach is extremely valuable considering the time-consuming character of field surveying [25]. In this regard, the *RMSE* values obtained in the statistical analysis confirm that 1 m resolution DEMs are reasonably better than DEMs generated at 2 m resolution using the PNOA-LiDAR data set. In this case, the IDW method achieved the best results applying both spatial resolutions (1 and 2 m), obtaining *RMSEs* of 37.10 cm and 40.60 cm, respectively. On the contrary, point to raster method presented the worst results, both at 1 m and 2 m resolution, with *RMSEs* of 50.90 and 63.00 cm, respectively. These low accuracy results match those shown in Table 3. The *ME* indicated a general overestimation of the elevation in all interpolation routines. This can be explained by the presence of systematic errors related with the filtering procedure. It is relatively frequent in ALS-derived DEMs of forest areas, where non-ground returns such as low vegetation and logs, are included as ground returns and subsequently in the DEM interpolation process, resulting in positive *MEs* [30,39]. The greatest range between minimum and maximum errors was obtained with the point to raster interpolator (1.90 m), indicating an exaggeration of the elevation errors. In addition, it should be stressed that vertical accuracy of GPS points is eight times better than the ALS data set. This fact, could also explain the higher *RMSE*, *MAE* and *ME* values obtained with the ground control points with respect to those obtained with ALS test data set, which is much more dependent of error acquisition and filtering process.

Table 3. Global statistics summarizing validation errors obtained with the ALS training dataset. Only the most accurate parameterization values for all the methods applied at both resolutions (1 and 2 m) are analyzed. *Mean Error (ME)*, *Mean Absolute Error (MAE)*, *Root Mean Square Error (RMSE)* and *Standard Deviation of residuals (SD)*. Interpolation methods are ranked from lowest to highest *RMSE*. $N = 1,896,468$.

Interpolation Method	Parameterization	Resolution (m)	Min Error (m)	Max Error (m)	Range (m)	SD (m)	ME (cm)	RMSE (cm)	MAE (cm)
TIN to raster	Applying natural neighbor interpolation to TIN triangles to obtain cell values.	1	−3.14	3.50	6.64	0.16	0.59	2.68	11.73
Natural neighbor	Not applicable.	1	−4.96	3.61	8.57	0.17	0.03	2.95	12.14
ANUDEM	Surface created with no drainage enforcement.	1	−3.06	3.31	6.37	0.17	−0.1	2.99	12.14
IDW	Power of 2 and variable search radius with 12 minimum points.	1	−2.77	3.81	6.58	0.19	−0.32	3.64	12.9
Kriging	Parameters determined fitting the semivariogram model. 4 sectors with an offset of 45° for the search neighborhood.	1	−4.42	4.44	8.86	0.20	−0.9	3.91	14.08
Point to raster	Not applicable.	1	−4.45	4.24	8.69	0.26	−1.29	6.64	18.25

Table 3. Cont.

Interpolation Method	Parameterization	Resolution (m)	Min Error (m)	Max Error (m)	Range (m)	SD (m)	ME (cm)	RMSE (cm)	MAE (cm)
Kriging	Parameters determined fitting the semivariogram model. 1 sector for the search neighborhood.	2	−2.71	4.23	6.94	0.23	0.48	5.25	16.25
ANUDEM	Surface created with no drainage enforcement.	2	−2.75	3.71	6.46	0.23	0.16	5.42	16.73
TIN to raster	Applying natural neighbor interpolation to TIN triangles to obtain cell values.	2	−4.14	3.29	7.43	0.23	0.86	5.48	16.94
Natural neighbor	Not applicable.	2	−3.26	3.70	6.96	0.23	0.86	5.52	16.94
IDW	Power of 2 and variable search radius with 12 minimum points.	2	−2.76	3.81	6.57	0.24	0.32	5.74	16.74
Point to raster	Not applicable.	2	−4.40	4.08	8.48	0.42	−2.73	17.67	29.54

Table 4. Global statistics summarizing validation errors using the GPS benchmarks. Only the most accurate parameterization values for all the methods applied at both resolutions (1 m and 2 m) are analyzed. *Mean Error (ME)*, *Mean Absolute Error (MAE)*, *Root Mean Square Error (RMSE)* and *Standard Deviation of residuals (SD)*. Interpolation methods are ranked from lowest to highest *RMSE*. $N = 55$.

Interpolation Method	Parameterization	Resolution (m)	Min Error (m)	Max Error (m)	Range (m)	SD (m)	ME (cm)	RMSE (cm)	MAE (cm)
IDW	Power of 1 and variable search radius with 24 minimum points.	1	−0.05	1.24	1.29	0.22	29.80	37.10	30.01
Kriging	Parameters determined fitting the semivariogram model. 1 sector for the search neighborhood.	1	−0.03	1.23	1.26	0.24	30.10	38.10	30.19
Natural neighbor	Not applicable.	1	−0.02	1.28	1.31	0.24	32.80	40.40	32.88
TIN to raster	Applying linear interpolation to TIN triangles to obtain cell values.	1	0.03	1.44	1.40	0.28	32.70	42.80	32.67
ANUDEM	Surface created with drainage enforcement.	1	−0.04	1.49	1.53	0.31	33.90	45.40	34.16
Point to raster	Not applicable.	1	−0.13	1.77	1.90	0.36	36.80	50.90	37.29
IDW	Power of 0.5 and variable search radius with 6 minimum points.	2	−0.29	1.25	1.55	0.28	29.70	40.60	31.11
ANUDEM	Surface created with drainage enforcement.	2	−0.12	1.27	1.39	0.28	32.40	42.70	33.79
Kriging	Parameters determined fitting the semivariogram model. 1 sector for the search neighborhood.	2	−0.03	1.58	1.61	0.30	32.50	44.10	32.67
Natural neighbor	Not applicable.	2	−0.05	1.86	1.90	0.32	33.20	46.00	33.37
TIN to raster	Applying linear interpolation to TIN triangles to obtain cell values.	2	−0.04	1.45	1.48	0.33	34.00	47.10	34.13
Point to raster	Not applicable.	2	−0.15	2.53	2.68	0.49	40.00	63.00	41.10

Figure 2 shows the results of the analysis of the influence of terrain slope, land cover, ground return density and canopy pulse penetration on the accuracy of the ALS-derived DEMs.

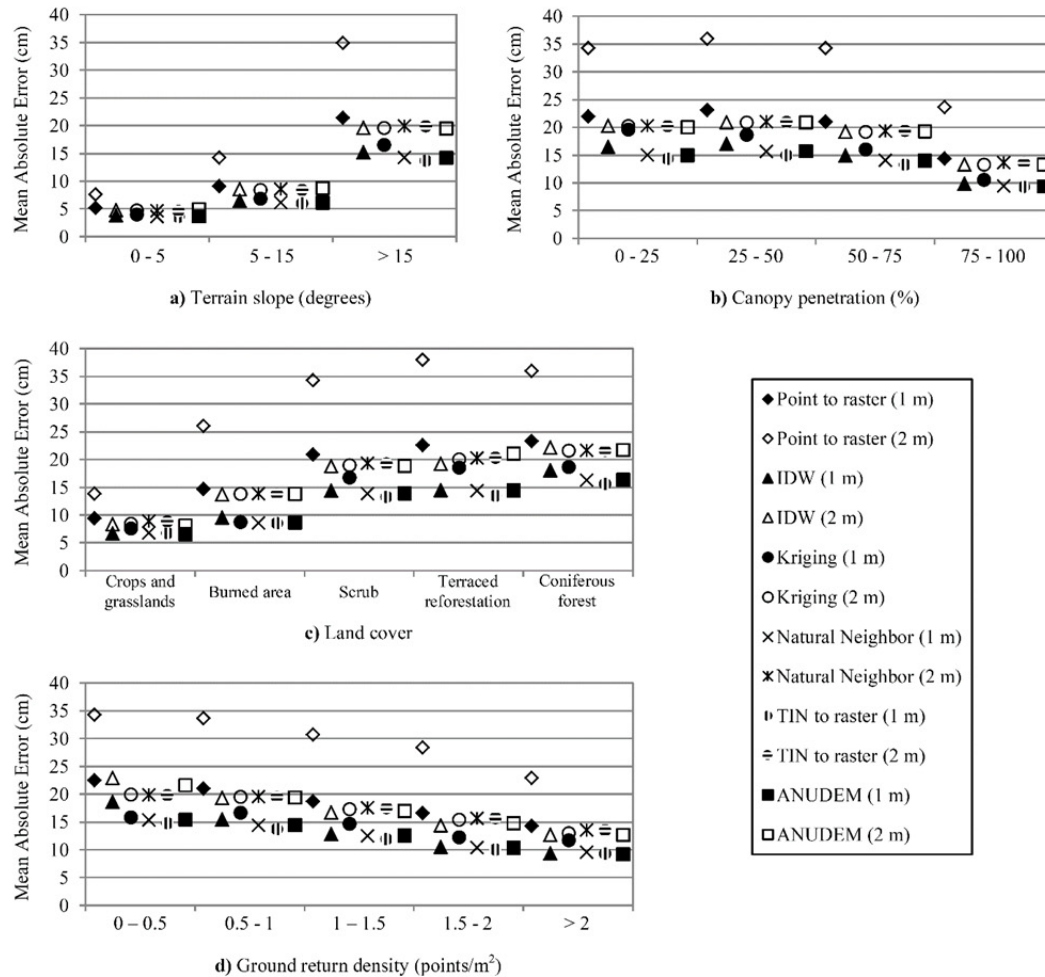


Figure 2. Effects of terrain slope (a); canopy pulse penetration (b); land cover (c) and (d) ground return density on MAE of interpolation.

In relation to the slope steepness, an increment from 0° to more than 15° increases the MAEs in all interpolation algorithms in more than 10 cm (Figure 2a). The finer spatial resolution DEMs presented higher accuracy than the coarser ones [22]. The 1 m surface created using TIN to raster method shows the lowest MAE across the range of slope classes.

On the contrary, MAE decreased with an increase in the penetration rate of laser pulses reaching the ground (Figure 2b). TIN to raster method performed the best when penetration rate was 0% to 25%, with a MAE of 14.32 cm, being also the most appropriate in all penetration rates. In this regard, among the interpolation routines tested, kriging presented the worst results at 1 m resolution. It should be noted that all methods, except kriging at 1 m resolution, experiment a slightly increase of approximately 1 cm in their MAEs from the 0%–25% penetration class to the 25%–50% category. This effect is prone to be due

to the error introduced by “Cover” command, inherent to the parameters chosen to map this variable for the analysis.

Land cover also affected interpolation accuracy (Figure 2c). The highest *MAEs* occurred in the coniferous forest class, which is characterized by the presence of Aleppo pine trees. In contrast, *MAEs* located in terraced reforestation were slightly lower, as a result of the gaps in the forest canopy allowing a larger proportion of ALS returns to reach the ground. In general, whether land cover structural complexity decreases, the error also decreases. As expected, the 2 m resolution DEMs present lower accuracy than 1 m resolution surfaces across all land cover classes. TIN to raster method was the best method to interpolate ALS data in presence of scrub and pine forests, as well as sprouted scrub, abandoned logs, stumps and woody debris typical of a burned area, presenting errors ranging from 8.45 to 15.60 cm. However, this method was slightly less suitable for crops and grasslands land covers than ANUDEM one, as the last presented a lower *MAE* of 6.50 cm, compared with the 6.69 cm of error of the TIN to raster method.

Figure 2d shows the effects of ALS ground return density on interpolation accuracy. The greatest *MAE* was lower than 23 cm and was produced in DEMs generated at 1 m resolution using the point to raster routine in areas with the lowest point densities, *i.e.*, between 0 and 0.5 points/m². Again, this method achieved the worst results, especially in 2 m resolution DEMs. In general, the rest of routines performed slightly better at 1 m resolution than at 2 m resolution. In addition, TIN to raster method presented the lowest *MAEs* with the exception of the ground return density range greater than 2 points/m². In this category, ANUDEM (*MAE* of 9.21 cm) and IDW (*MAE* of 9.22 cm) had a better performance.

Finally, the results obtained after performing the Kruskal Wallis test show that the differences between the errors obtained applying different interpolations routines are statistically significant. In summary, DEMs generated at 1 m resolution present a higher accuracy than those of 2 m resolution and the TIN to raster method seems to be the most suitable one to interpolate ALS data of low point densities (<0.5 points/m²) in Mediterranean forested environments characterized by a variable slope steepness and a relatively complex landscape.

3.3. Error Prediction

In order to ensure consistency in the analysis, the DEM generated at 1 m resolution using TIN to raster methodology was used in the final accuracy assessment. The error prediction map created from the CART analysis is presented in Figures 3 and 4. Similar to Bater and Coops [22], the CART analysis indicated that the most important predictor variables in interpolation error were terrain slope and point density, but also land cover, which determines the amount of returns reaching the ground.

As expected, areas with a combination of high slope (>15°) and low ground points density (≤0.30 points/m²) were the most prone to interpolation error. However, in terrain with slope steepness lower than 15°, prediction uncertainty was very low. This method is advantageous where the combined effects of the two predictor variables are less intuitive, for instance, high slope (>15°) and high point density (>1.07 points/m²). In the end, using prediction uncertainty maps may help in the detection of potential problems with ALS-derived vegetation height estimates in those areas where the DEM surface is uncertain [22].

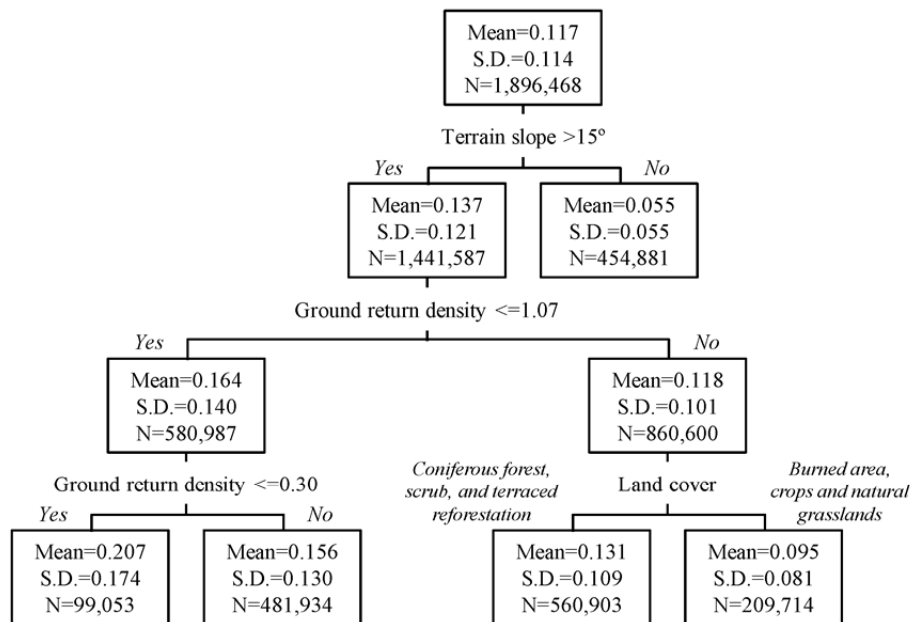


Figure 3. Classification tree resulting from CART analysis of absolute errors for a 1 m resolution DEM created using TIN to raster interpolation. Each node (square) is labeled with *average absolute error (Mean)*, *standard deviation (S.D.)* and the *number (N)* of points in that group. The model is read from top down until terminal nodes predicting the vertical error from selected variables appear.

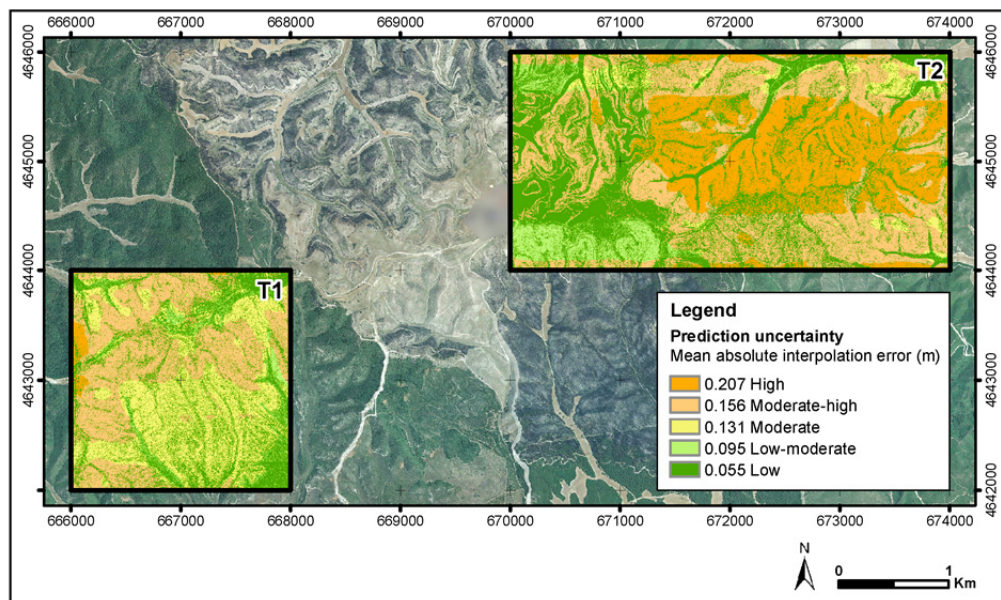


Figure 4. CART-derived prediction uncertainty map over the high spatial resolution orthophotography (PNOA-2009) used as backdrop for 1 m TIN to raster DEM. For this surface, slope, ground return density and land cover were the best predictors of interpolation error.

Figure 4 shows the classification of the study area into categories of prediction uncertainty. The eastern half of T2 showed higher uncertainty, particularly compared to the southwest of this test site. Good results are shown in the flat-bottom valleys occupied by field crops and in the burned area, except in zones occupied by sprouting shrub vegetation and abandoned logs. Furthermore flight-overlaps strips can be observed in T2, which implies more point density and therefore less uncertainty in DEM surface. In general, T1 test site shows moderate error uncertainty, mainly due to the hilly relief, the high point density in the area, but also due to the presence of the pine forests and the shrub vegetation cover. Figure 4 corroborates that topographic gradient and a low point density are the main factors in DEM error. In fact, vertical errors were observed in areas where the local variability in the terrain (e.g., topographic slope) was large and when ALS point count was low (*i.e.*, ≤ 0.30 points/m²).

4. Discussion

The aim of this research was comparing the performance of different interpolation techniques to derive gridded DEMs in a Mediterranean forested region in order to normalize the PNOA-LiDAR data. Our results establish the first baseline for potential users of low density point clouds, in absence of information describing the suitability of interpolation parameters in areas occupied by Aleppo pine forest mixed with evergreen shrub. This is a contribution to other researches like those of Rees [65], Lloyd and Atkinson [12] and Mitášová *et al.* [66].

Bater and Coops [22], and Lloyd and Atkinson [67] pointed out in their research that no interpolation method is universally superior since ground return spacing, raster pixel size, the complexity of terrain morphology, and the assumptions of a given interpolator affect the ability of interpolation routines to generate accurate DEMs. In our study, the *RMSE* obtained in the DEM validation using the ALS test data set, varied with different methods and resolutions. However, our research confirmed that natural neighbor, IDW, kriging, ANUDEM, and TIN to raster do not differ greatly in terms of their global *RMSEs* and *MAEs* for resolutions 1 and 2 m (see Tables 2 and 3). The ANUDEM and IDW interpolators were the most conservative routines obtaining the lowest range of error. In this sense, Bater and Coops [22], found more conservative the linear and natural neighbor methods. In our case, at 1 m resolution, TIN to raster and natural neighbor, the two best interpolators, had an *RMSE* of 2.68 and 2.95 cm, respectively, while point to raster, the worst interpolator, resulted in an *RMSE* of 6.64 cm. This indicates that the errors produced by interpolators are as significant as the measurement errors and should be considered when generating high quality DEMs from ALS data [26]. On the other hand, *MEs* showed that the two best interpolators slightly overestimated the ground elevation. This effect is frequently encountered when working with ALS data in forested areas where is usual to find a positive bias as the point cloud is misclassified due to the presence of dense low lying vegetation under the tree canopy [5,32,39,68]. This overestimation is attributed to the reduced number of ALS ground points used to interpolate each grid centroid [36] but also to the use of vegetation point as ground points to generate the DEMs. Despite this fact, overestimations were limited in our research to 0.59 cm and 0.03 cm, using TIN to raster and natural neighbor methodologies, respectively. This would likely have little impact on attributes of vegetation structure derived from these DEMs. In general the *ME* values around zero obtained in all tested routines to generate DEMs, suggest unbiased predictions according to Gallay *et al.* [64] and Bater and Coops [22].

The statistics fall below a centimeter level with the exception of point to raster interpolation, where the predictions were systematically underestimated by less than 3 cm considering the *MEs* values.

The supplemental validation of DEMs with the 55 ground surveyed measurements with a high precision GPS provided insight into the absolute accuracy of the bare-earth surfaces. Our results fall within the typical *RMSE* values reported by other empirical studies, which ranged from 0.14 to 1.50 m, depending on the operational aspects of ALS and environmental conditions [6,38].

The *RMSE* values showed that 1 m resolution DEMs are reasonably better than 2 m resolution DEMs. As confirmed Gongga-Saholiariliva *et al.* [27], it was shown that the larger the grid-cell size, the lower the accuracy in DEMs. However, the higher *MEs* obtained in general, ranging from 40.00 to 29.70 cm with the GPS ground control points suggest the necessity of a thorough analysis, using more control points for a higher level of confidence in the validation results. However, it was not possible to improve the number of checkpoints, as according to Liu [25] the collection with GPS of a large number of high-accuracy checkpoints was a time-consuming task, leading to an increase of the costs of the study.

The results of ALS-derived DEMs validated with the GPS benchmark tended to overestimate the reference ground elevation [6]. The IDW interpolation method presented the best results, *RMSE* of 37.10 cm, using a power of 1 and a variable search radius with 24 minimum points. As pointed out Gallay *et al.* [64], the IDW is an approachable method in proprietary as well as open source GIS or statistical software and it is relatively easy to parameterize. However, Renslow [2] underlined that in some ALS data sets where the point density varies widely, this method can be a challenge since different densities of points often dictate different parameters for best results. In this regard, the amount of nearest neighbors used in IDW interpolation may seem high. However it is known that a very small number of points are prone to cause artifacts in the DEM, at least in certain cases based on the interaction with the spatial pattern of the point cloud.

Additionally, computation time should be considered when choosing the appropriate interpolation method, although the absolute computation time may change under different computation conditions, such as the computer's CPU, available memory, and software used [26]. The computations for this research ran under a Windows™ server with Intel® Core™ i5 3.10 GHz processors and 8.00 GB memory using ESRI ArcGIS 10.1 (ESRI, Redlands, CA, USA) and FUSION LDV 3.30 [43]. Natural neighbor, point to raster and TIN to raster prove to be simple and fast methods (60, 120 and 198 seconds to generate each DEM, respectively), while IDW and kriging are moderate in computation time (240 and 540 seconds) compared with the ANUDEM (1140 seconds), which is the slowest one. These results are in line with those obtained by Guo *et al.* [26] where the simplest methods have the best processing time.

In terms of the accuracy level considering the pixel size, we have demonstrated, like Bilskie and Hagen [28] that *MAE* and *RMSE* generally increase along with larger DEM grid cells. As the DEM becomes coarser, it is unable to describe sub-scale undulations of the ground surface that are better represented by higher resolution DEMs. Nevertheless, the findings of Rees [65], Liu *et al.* [11] and Smith *et al.* [69] show that the choice of interpolation method is less influential when a surface is interpolated to coarser resolutions than the resolution of the input data. In any case, our research suggests that PNOA-LiDAR-derived DEMs with pixel size similar to point density, 1 m, achieved very good results. In this context, Behan [70] quantified the error within models produced from different

interpolation algorithms and obtained the highest accuracy in surfaces created using cell sizes with spacing analogous to the original points.

Ultimately, TIN to raster interpolation is usually preferred due to the overall simplicity nature of its performance, as it has no adjustable parameters, so no user-tunable variances are introduced, and to its efficiency in processing [2]. In this sense, TIN to raster interpolation is also the best option for interpolating the ground returns of PNOA-LiDAR data in a forested Mediterranean environment since it presents a consistent accuracy and relative conservative predictions.

On the other hand, analysis of the effects of terrain slope, land cover, ALS ground point density and pulse penetration on DEMs accuracy showed that all factors influence *MAEs*. Although it is known that ALS-derived DEMs are less sensitive to terrain slope than those derived from digital photogrammetry [6,38], topographic gradients are a significant factor in DEM error as corroborated by our research and others such as Hodgson and Bresnahan [38], Su *et al.* [32], Gallay *et al.* [64] and Bater and Coops [22]. As Aguilar *et al.* [6] suggested, ALS planimetric error may be relatively high (up to 0.30 m for PNOA-LiDAR mission) and also may be directly translated to vertical errors on sloping surfaces. In our research, double *MAE* values were obtained in areas with moderate slope steepness (from 5° to 15°) in comparison with areas of low slope steepness (from 0° to 5°) in 1 m resolution DEM. Similar patterns were obtained by Hodgson and Bresnahan [38], who observed elevation error in steeper slopes (about 25°) twice of those observed on low slopes (e.g., 15°). Like Bater and Coops [22], and increase in slope steepness from 0° to more than 15° caused decimeter-level increases in *MAEs* in all interpolation routines, especially in point to raster method. Except for this method, the rest of interpolation methods were similar in their accuracies, although the kriging and IDW routines appeared to be more sensitive to changes in slope.

Hodgson and Bresnahan [38], Aguilar and Mills [39] and Aguilar *et al.* [6], also indicated the influence of land-cover in the accuracy of DEMs. In general, as the structural complexity of the land cover decreased, the *MAEs* obtained in the DEMs generated also diminished. Our results confirm, as Bater and Coops [22] and Hodgson and Bresnahan [38] pointed out in their studies, that higher *MAEs* (ranging from 15 to 40 cm) occurred in areas with tall canopy vegetation, covered by dense coniferous forest. However, Hodgson and Bresnahan [38] found the largest *RMSE* in areas covered by scrub. In connection with this, the presence of vegetation can limit ground detection, due to a decrease in the canopy pulse penetration. Nevertheless, this is not only a deficiency of ALS data, but also in DEMs created from stereo photogrammetry, radar, or ground surveying, where the accuracy and reliability of the surface generated is usually lower in vegetated areas than in open areas [2,25]. Laser energy often fails to penetrate a dense vegetation canopy resulting in last returns that are well above the true ground surface [13]. Our study confirms that the TIN to raster method was the most appropriate for the different penetration rates, as it had the lowest *MAEs* in the ground point density categories, ranging from 14.98 to 9.24 cm, with the exception of areas with ground penetration densities greater than 2 points/m², where ANUDEM presented one centimeter more in accuracy. This support the conclusion of Hu *et al.* [71], who confirmed that increasing sampling density implies a decrease in interpolation error.

Finally, the CART analysis indicated that topographic variability and sampling density have significant influence on the accuracy of ALS-derived DEMs but also the land cover, which determines the amount of returns reaching the ground. As pointed out by Guo *et al.* [26], whether the complexity of the terrain increases, the uncertainty in the derived DEM also increases. In this regard, our results are

similar to previous studies conducted by Bater and Coops [22] and Hodgson and Bresnahan [38], who indicated that the pattern of highest magnitude error was observed to occur in the areas of greater surface roughness. In addition, Aguilar *et al.* [72] found that morphology has the greatest influence on DEM quality, followed by the sampling density and interpolation method. Similar to Aguilar *et al.* [6], the error in ALS-derived DEMs is not very sensitive to change in point density in areas of low average slope as can be seen in Figure 3. According to Aguilar *et al.* [6], it should be noted that the total vertical error of the MDE can be disaggregated into three main components: (i) the error from ALS data capture; (ii) the error due to filtering method; and (iii) the error of interpolation method and gridding.

In order to improve the results presented in this study, we consider important for future investigation providing a replicated analysis of density reduction to suggest possible strategies for reducing ALS data sets sizes as Anderson *et al.* [36] performed. As data sets become more widely available across larger areas and, subsequently, data set size will often be prohibitively large, the computational requirements for handling such data will become an even greater issue [36]. With a reduction in data, a more usable and operationally sized elevation data set will be possible, increasing the efficiency in terms of storage and manipulation [23]. Moreover, an empirical model to quantify the relationship between DEM and the influencing factors analyzed in this paper could be developed following the trend proposed by Aguilar *et al.* [29] for estimating global and absolute accuracy and predicting the error budget after applying ALS data filtering and gridding processes.

5. Conclusions

The selection of an appropriate interpolation method and spatial resolution becomes an important decision in DEM generation. This paper focuses on the assessment of six interpolation methods to generate an optimal DEM in order to normalize the ALS data captured by the PNOA-LiDAR mission to estimate vegetation structural metrics in a Mediterranean forested landscape. The interpolation methods analyzed include natural neighbor, IDW, ordinary kriging, ANUDEM, TIN to raster, and point to raster approaches. A collection of DEMs was generated with a spatial resolution of 1 and 2 m, according to the ground point density. Then, the accuracy of the ALS-derived DEMs was assessed with a test sample of ALS points and complementary with an independent reference set of 55 ground control points collected randomly on foot with a high precision GPS. The results of the validation using the ALS test samples of points showed a higher accuracy of DEMs created with the TIN to raster interpolation method at 1 m resolution grid. On the contrary, kriging interpolation was the best at 2 m resolution DEM. Poor accuracy was achieved with point to raster routine, which is considered the simplest means of converting point data to a raster surface. The high *RMSEs* obtained in general with the GPS control points, showed that the IDW present the lowest *RMSE* both, at 1 and 2 m resolution. Overall, the results confirmed that 1 m resolution DEMs present a higher accuracy than 2 m resolution ones. Additionally, with the purpose of examining the effect of terrain slope, land cover, ground point density and pulse penetration on interpolation error, the study area was stratified by these variables. Based on the error statistics computed, we concluded that the TIN to raster interpolation was the optimal solution in any terrain slope steepness, in areas with low point densities (below 0.5 points/m²) and complex land cover, such as scrub and pine forests, as well as sprouted scrub, abandoned logs, stumps and woody debris, typical of a burned area. Finally, the CART analysis allowed us to conclude that areas with a combination of high slope

steepness (above 15°) and low point density (below 0.3 points/m²) were the most prone to present high interpolation errors.

Acknowledgments

This research has been financed by the Government of Aragón, Department of Science, Technology and University (FPI Grant BOA 30, 11/02/2011) and supported by the Research Projects Ref. No 2013-04 of Centro Universitario de la Defensa de Zaragoza, and Ref. UZCUD2014-HUM-01 of University of Zaragoza and Centro Universitario de la Defensa de Zaragoza. The ALS data were provided by the National Center for Geographic Information of Spain. The authors are grateful to the Training Center (CENAD) “San Gregorio” for assistance in the field.

Author Contributions

María Teresa Lamelas had the original idea for the study. Antonio L. Montealegre and María Teresa Lamelas conducted fieldwork campaign. Antonio L. Montealegre, María Teresa Lamelas and Juan de la Riva developed the methodology and performed the analysis and Antonio L. Montealegre wrote the manuscript, incorporating suggestions from all co-authors, who approved the final manuscript.

Conflicts of Interest

The authors declare no conflict of interest.

References

1. Huising, E.J.; Gomes Pereira, L.M. Errors and accuracy estimates of laser data acquired by various laser scanning systems for topographic applications. *ISPRS J. Photogramm. Remote Sens.* **1998**, *53*, 245–261.
2. Renslow, M. *Manual of Airborne Topographic LiDAR*; The American Society for Photogrammetry and Remote Sensing: Bethesda, MD, USA, 2013.
3. Baltsavias, E.P. A comparison between photogrammetry and laser scanning. *ISPRS J. Photogramm. Remote Sens.* **1999**, *54*, 83–94.
4. Razak, K.A.; Straatsma, M.W.; van Westen, C.J.; Malet, J.P.; de Jong, S.M. Airborne laser scanning of forested landslides characterization: Terrain model quality and visualization. *Geomorphology* **2011**, *126*, 186–200.
5. Kraus, K.; Pfeifer, N. Determination of terrain models in wooded areas with airborne laser scanner data. *ISPRS J. Photogramm. Remote Sens.* **1998**, *53*, 193–203.
6. Aguilar, F.J.; Mills, J.P.; Delgado, J.; Aguilar, M.A.; Negreiros, J.G.; Pérez, J.L. Modelling vertical error in LiDAR-derived digital elevation models. *ISPRS J. Photogramm. Remote Sens.* **2010**, *65*, 103–110.
7. Kraus, K.; Pfeifer, N. Advanced DTM generation from LiDAR data. *Int. Arch. Photogramm. Remote Sens. Spat. Inf. Sci.* **2001**, *34*, 23–30.
8. Lohr, U. Digital elevation models by laser scanning. *Photogramm. Rec.* **1998**, *16*, 105–109.

9. Wehr, A.; Lohr, U. Airborne laser scanning: An introduction and overview. *ISPRS J. Photogramm. Remote Sens.* **1999**, *54*, 68–82.
10. Lefsky, M.A.; Cohen, W.B.; Parker, G.G.; Harding, D.J. LiDAR remote sensing for ecosystem studies. *BioScience* **2002**, 19–30.
11. Liu, X.; Zhang, Z.; Peterson, J.; Chandra, S. LiDAR-derived high quality ground control information and DEM for image orthorectification. *GeoInformatica* **2007**, *11*, 37–53.
12. Lloyd, C.D.; Atkinson, P.M. Deriving ground surface digital elevation models from LiDAR data with geostatistics. *Int. J. Geogr. Inf. Sci.* **2006**, *20*, 535–563.
13. Vosselman, G.; Maas, H.-G. *Airborne and Terrestrial Laser Scanning*; Whittles Publishing: Dunbeath, UK, 2010.
14. Næsset, E. Determination of mean tree height of forest stands using airborne laser scanner data. *ISPRS J. Photogramm. Remote Sens.* **1997**, *52*, 49–56.
15. Næsset, E. Estimating timber volume of forest stands using airborne laser scanner data. *Remote Sens. Environ.* **1997**, *61*, 246–253.
16. Means, J.E.; Acker, S.A.; Fitt, B.J.; Renslow, M.; Emerson, L.; Hendrix, C.J. Predicting forest stand characteristics with airborne LiDAR. *Photogramm. Eng. Remote Sens.* **2000**, *66*, 1367–1371.
17. Lefsky, M.A.; Cohen, W.B.; Acker, S.A.; Parker, G.G.; Spies, T.A.; Harding, D. LiDAR remote sensing of the canopy structure and biophysical properties of douglas-fir western hemlock forests. *Remote Sens. Environ.* **1999**, *70*, 339–361.
18. Dubayah, R.O.; Drake, J.B. LiDAR Remote sensing for forestry. *J. For.* **2000**, *98*, 44–46.
19. Lim, K.; Treitz, P.; Wulder, M.; St-Onge, B.; Flood, M. LiDAR remote sensing of forest structure. *Prog. Phys. Geogr.* **2003**, *27*, 88–106.
20. Riaño, D.; Meier, E.; Allgöwer, B.; Chuvieco, E.; Ustin, S.L. Modeling airborne laser scanning data for the spatial generation of critical forest parameters in fire behavior modeling. *Remote Sens. Environ.* **2003**, *86*, 177–186.
21. Hall, S.A.; Burke, I.C.; Box, D.O.; Kaufmann, M.R.; Stoker, J.M. Estimating stand structure using discrete-return LiDAR: An example from low density, fire prone ponderosa pine forests. *For. Ecol. Manag.* **2005**, *208*, 189–209.
22. Bater, C.W.; Coops, N.C. Evaluating error associated with LiDAR-derived DEM interpolation. *Comput. Geosci.* **2009**, *35*, 289–300.
23. Vaze, J.; Teng, J.; Spencer, G. Impact of DEM accuracy and resolution on topographic indices. *Environ. Model. Softw.* **2010**, *25*, 1086–1098.
24. Sithole, G.; Vosselman, G. Experimental comparison of filter algorithms for bare-Earth extraction from airborne laser scanning point clouds. *ISPRS J. Photogramm. Remote Sens.* **2004**, *59*, 85–101.
25. Liu, X. Accuracy assessment of LiDAR elevation data using survey marks. *Surv. Rev.* **2011**, *43*, 80–93.
26. Guo, Q.; Li, W.; Yu, H.; Alvarez, O. Effects of topographic variability and LiDAR sampling density on several DEM interpolation methods. *Photogramm. Eng. Remote Sens.* **2010**, *76*, 701–712.
27. Gonga-Saholiariliva, N.; Gunnell, Y.; Petit, C.; Mering, C. Techniques for quantifying the accuracy of gridded elevation models and for mapping uncertainty in digital terrain analysis. *Progr. Phys. Geogr.* **2011**, *35*, 739–764.

28. Bilskie, M.V.; Hagen, S.C. Topographic accuracy assessment of bare earth LiDAR-derived unstructured meshes. *Adv. Water Resour.* **2013**, *52*, 165–177.
29. Aguilar, F.J.; Aguilar, M.A.; Agüera, F.; Sánchez, J. The accuracy of grid digital elevation models linearly constructed from scattered sample data. *Int. J. Geogr. Inf. Sci.* **2006**, *20*, 169–192.
30. Clark, M.L.; Clark, D.B.; Roberts, D.A. Small-footprint LiDAR estimation of sub-canopy elevation and tree height in a tropical rain forest landscape. *Remote Sens. Environ.* **2004**, *91*, 68–89.
31. Hopkinson, C.C.; Chasmer, L.E.; Zsigovics, G.; Creed, I.F.; Sitar, M.; Treitz, P.; Maher, R.V. Errors in LiDAR ground elevation and wetland vegetation height estimates. *Proc. Int. Arch. Photogramm. Remote Sens. Spat. Inf. Sci.* **2004**, *36*, 108–113.
32. Su, J.; Bork, E. Influence of vegetation, slope and LiDAR sampling angle on DEM accuracy. *Photogramm. Eng. Remote Sens.* **2006**, *72*, 1265–1274.
33. Desmet, P.J.J. Effects of interpolation errors on the analysis of DEMs. *Earth Surf. Process. Landf.* **1997**, *22*, 563–580.
34. Gong, J.; Li, Z.; Zhu, Q.; Sui, H.; Zhou, Y. Effects of various factors on the accuracy of DEMs: An intensive experimental investigation. *Photogramm. Eng. Remote Sens.* **2000**, *66*, 1113–1117.
35. Schoorl, J.M.; Sonneveld, M.P.W.; Veldkamp, A. Three-dimensional landscape process modelling: The effect of DEM resolution. *Earth Surf. Process. Landf.* **2000**, *25*, 1025–1034.
36. Anderson, E.S.; Thompson, J.A.; Crouse, D.A.; Austin, R.E. Horizontal resolution and data density effects on remotely sensed LIDAR-based DEM. *Geoderma* **2006**, *132*, 406–415.
37. Fisher, P.F.; Tate, N.J. Causes and consequences of error in digital elevation models. *Prog. Phys. Geogr.* **2006**, *30*, 467–489.
38. Hodgson, M.E.; Bresnahan, P. Accuracy of airborne LiDAR-derived elevation. *Photogramm. Eng. Remote Sens.* **2004**, *70*, 331–339.
39. Aguilar, F.J.; Mills, J.P. Accuracy assessment of LiDAR-derived digital elevation models. *Photogramm. Rec.* **2008**, *23*, 148–169.
40. Spanish National Plan for Aerial Orthophotography (PNOA). Available online: <http://pnoa.ign.es/presentacion> (accessed on 2 July 2015).
41. 3D Elevation Program (3DEP) of the U.S. Geological Survey National Geospatial Program (NGP). Available online: <http://nationalmap.gov/3DEP/> (accessed on 13 June 2015).
42. Gatzliolis, D.; Fried, J.S.; Monleon, V.S. Challenges to estimating tree height via LiDAR in closed-canopy forests: A parable from western Oregon. *For. Sci.* **2010**, *56*, 139–155.
43. McGaughey, R. *FUSION/LDV: Software for LIDAR Data Analysis and Visualization*; US Department of Agriculture, Forest Service, Pacific Northwest Research Station: Seattle, WA, USA, 2009.
44. Vicente-Serrano, S.M.; Lasanta, T.; Gracia, C. Aridification determines changes in forest growth in *Pinus halepensis* forests under semiarid Mediterranean climate conditions. *Agric. For. Meteorol.* **2010**, *150*, 614–628.
45. Notivol, E.; Cabanillas, A.; González, R.; Revuelta, C. Caracterización de masas naturales de pino carrasco (*Pinus halepensis* Mill.) en la Depresión del Ebro. In Proceedings of the La Ciencia Forestal: Respuestas para la Sostenibilidad. 4º Congreso Forestal Español, Zaragoza, Spain, 26–30 September 2005; p. 7.

46. Tanase, M.; de la Riva, J.; Pérez-Cabello, F. Estimating burn severity at the regional level using optically based indices. *Can. J. For. Res.* **2011**, *41*, 863–872.
47. Tanase, M.; de la Riva, J.; Santoro, M.; Pérez-Cabello, F.; Kasischke, E. Sensitivity of SAR data to post-fire forest regrowth in Mediterranean and boreal forests. *Remote Sens. Environ.* **2011**, *115*, 2075–2085.
48. Chuvieco, E. *Earth Observation of Wildland Fires in Mediterranean Ecosystems*; Springer-Verlag: Berlin/Heidelberg, Germany, 2009.
49. Evans, J.S.; Hudak, A.T. A multiscale curvature algorithm for classifying discrete return LiDAR in forested environments. *IEEE Trans. Geosci. Remote Sens.* **2007**, *45*, 1029–1038.
50. Montealegre, A.L.; Lamelas, M.T.; de la Riva, J. A comparison of open source LiDAR filtering algorithms in a Mediterranean forest environment. *IEEE Trans. Sel. Top. Appl. Earth Observ. Remote Sens.* **2015**, *2015*, doi:10.1109/JSTARS.2015.2436974.
51. Sibson, R. A brief description of natural neighbour interpolation. In *Interpreting Multivariate Data*; Barnett, V., Ed.; John Wiley & Sons: New York, NY, USA, 1981; pp. 21–36.
52. Sambridge, M.; Braun, J.; McQueen, H. Geophysical parametrization and interpolation of irregular data using natural neighbours. *Geophys. J. Int.* **1995**, *122*, 837–857.
53. Hutchinson, M.F. A new procedure for gridding elevation and stream line data with automatic removal of spurious pits. *J. Hydrol.* **1989**, *106*, 211–232.
54. Cressie, N. Spatial prediction and ordinary kriging. *Math. Geol.* **1988**, *20*, 405–421.
55. Caruso, C.; Quarta, F. Interpolation methods comparison. *Comput. Math. Appl.* **1998**, *35*, 109–126.
56. Zhang, X.F.; van Eijkeren, J.C.H.; Heemink, A.W. On the weighted least-squares method for fitting a semivariogram model. *Comput. Geosci.* **1995**, *21*, 605–608.
57. Chu, H.-J.; Chen, R.-A.; Tseng, Y.-H.; Wang, C.-K. Identifying LiDAR sample uncertainty on terrain features from DEM simulation. *Geomorphology* **2014**, *204*, 325–333.
58. Breiman, L.; Friedman, J.H.; Olshen, R.A.; Stone, C.J. *Classification and Regression Trees*; Wadsworth International Group: Pacific Grove, CA, USA, 1984.
59. Zheng, H.; Chen, L.; Han, X.; Zhao, X.; Ma, Y. Classification and regression tree (CART) for analysis of soybean yield variability among fields in Northeast China: The importance of phosphorus application rates under drought conditions. *Agric. Ecosyst. Environ.* **2009**, *132*, 98–105.
60. Tayyebi, A.; Pijanowski, B.C. Modeling multiple land use changes using ANN, CART and MARS: Comparing tradeoffs in goodness of fit and explanatory power of data mining tools. *Int. J. Appl. Earth Observ. Geoinf.* **2014**, *28*, 102–116.
61. Lawrence, R.L.; Wright, A. Rule-based classification systems using classification and regression tree (CART) analysis. *Photogramm. Eng. Remote Sens.* **2001**, *67*, 1137–1142.
62. Brown de Colstoun, E.C.; Story, M.H.; Thompson, C.; Comisso, K.; Smith, T.G.; Irons, J.R. National Park vegetation mapping using multitemporal Landsat 7 data and a decision tree classifier. *Remote Sens. Environ.* **2003**, *85*, 316–327.
63. Hladik, C.; Alber, M. Classification of salt marsh vegetation using edaphic and remote sensing-derived variables. *Estuar. Coast. Shelf Sci.* **2014**, *141*, 47–57.
64. Gallay, M.; Lloyd, C.; McKinley, J. Optimal interpolation of airborne laser scanner data for fine-scale DEM validation purposes. In Proceedings of Symposium GIS Ostrava 2012—Surface models for Geosciences, Ostrava, Czech Republic, 23–25 January 2012.

65. Rees, W.G. The accuracy of Digital Elevation Models interpolated to higher resolutions. *Int. J. Remote Sens.* **2000**, *21*, doi:10.1080/014311600210957.
66. Mitášová, H.; Mitáš, L.; Harmon, R.S. Simultaneous spline approximation and topographic analysis for ALS elevation data in open source GIS. *IEEE Geosci. Remote Sens. Lett.* **2005**, *2*, 375–379.
67. Lloyd, C.D.; Atkinson, P.M. Deriving DSMs from LiDAR data with kriging. *Int. J. Remote Sens.* **2002**, *23*, 2519–2524.
68. Goodwin, N.R.; Coops, N.C.; Culvenor, D.S. Assessment of forest structure with airborne LiDAR and the effects of platform altitude. *Remote Sens. Environ.* **2006**, *103*, 140–152.
69. Smith, S.L.; Holland, D.A.; Longley, P.A. Quantifying interpolation errors in urban airborne laser scanning models. *Geogr. Anal.* **2005**, *37*, 200–224.
70. Behan, A. On the matching accuracy of rasterized scanning laser altimeter data. *Int. Arch. Photogramm. Remote Sens.* **2000**, *33*, 75–82.
71. Hu, P.; Liu, X.; Hu, H. Accuracy assessment of digital elevation models based on approximation theory. *Photogramm. Eng. Remote Sens.* **2009**, *75*, 49–56.
72. Aguilar, F.J.; Agüera, F.; Aguilar, M.A.; Carvajal, F. Effects of terrain morphology, sampling density, and interpolation methods on grid DEM accuracy. *Photogramm. Eng. Remote Sens.* **2005**, *71*, 805–816.

© 2015 by the authors; licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution license (<http://creativecommons.org/licenses/by/4.0/>).

3.3 Estimación de variables dasométricas en masas de *P. halepensis* Mill.

RESUMEN

La teledetección activa con sensores LiDAR aporta una nueva perspectiva a los inventarios forestales al ofrecer de forma directa información 3-D del territorio, tanto de su topografía como de la estructura vegetal existente. Dada la disponibilidad de datos LiDAR de baja resolución espacial para todo el territorio español dentro del PNOA, en este artículo se evalúa su idoneidad para estimar variables dasométricas en masas regulares de *P. halepensis* Mill., localizadas en el entorno del Centro de Adiestramiento “San Gregorio” (Zaragoza).

Las nubes de puntos han sido filtradas y procesadas con los software MCC v.2.1 y FUSION LDV v.3.30, respectivamente. Posteriormente, se ha realizado un análisis multivariante de regresión lineal, relacionando las variables dasométricas de interés, estimadas en 45 parcelas de campo, con una colección de variables independientes extraídas de la nube de puntos. Los modelos han sido validados mediante la técnica de validación cruzada dejando un dato fuera (LOOCV).

Los coeficientes de determinación de los modelos de estimación han sido: 0,87 para la altura media, 0,84 para el diámetro cuadrático medio, 0,89 para el área basimétrica, 0,89 para el volumen con corteza, 0,48 para la densidad de pies y 0,69 para el ratio de copa. Las variables derivadas de los datos LiDAR introducidas en los modelos incluyeron estadísticos relacionados directamente con la altura del dosel, con la variabilidad, dispersión y forma de la distribución de las alturas de los puntos, y con la densidad y cobertura del dosel arbóreo.

Use of low point density ALS data to estimate stand-level structural variables in Mediterranean Aleppo pine forest

A. L. Montealegre^{1,2*}, M. T. Lamelas^{2,3}, J. de la Riva^{1,2}, A. García-Martín^{2,3} and F. Escribano³

¹Departamento de Geografía y O.T., Universidad de Zaragoza, Pedro Cerbuna 12, 50009 Zaragoza, Spain

²Grupo GEOFOREST-IUCA, Universidad de Zaragoza, Pedro Cerbuna 12, 50009 Zaragoza, Spain

³Centro Universitario de la Defensa de Zaragoza, Academia General Militar, Ctra. de Huesca s/n, 50090 Zaragoza, Spain

*Corresponding author. E-mail: monteale@unizar.es.

Received 5 July 2015

Little research has focussed on the use of low point density airborne laser scanning (ALS) technology in Mediterranean Aleppo pine (*Pinus halepensis* Mill.) forest to derive key stand characteristics for forestry. The objective of this research was to develop models of Aleppo pine structural variables from ALS point clouds with a point density of 1 point m⁻² and field information, in an area located in north-eastern Spain. Field stand metrics were obtained within 45 plots, 30 m in diameter. A multiple regression analysis was performed in order to develop the models for mean height, squared mean diameter, basal area, timber volume, stand density and crown ratio, which led to the following *R*² values: 0.87, 0.84, 0.89, 0.89, 0.48 and 0.69, respectively. ALS-derived metrics included in the models were maximum height, the 99th, 95th and 50th canopy height percentiles, the skewness, the kurtosis, the standard deviation, the coefficient of variation and the percentages of all and first returns from above 1 m of the ground surface. With the exception of stand density, the final models are appropriate to be used in forest management operations in Mediterranean Aleppo pine forest. This study demonstrates the usefulness of low point density ALS data to accurately estimate key structural parameters of Aleppo pine forest using an area-based approach.

Keywords: airborne laser scanning, LiDAR, forest inventory, Aleppo pine, multiple regression analysis, Mediterranean environment

Introduction

During the last decades, operational collection of information on the location and extent of many relevant characteristics of forest ecosystems by means of pure ground-based field inventories has been revolutionized by the development of remote sensing sensors (Maltamo *et al.*, 2014). Although optical and radar remote sensing have been widely used to map forest structural attributes and biophysical parameters (Le Toan *et al.*, 1992; Chuvieco, 2009; Tanase *et al.*, 2011; Leboeuf *et al.*, 2012), Light Detection and Ranging (LiDAR) has emerged as one of the very promising techniques for forest attribute estimation at several scales and applications (Lefsky *et al.*, 2002a; Vosselman and Maas, 2010; Maltamo *et al.*, 2014). The spatial position of all the recorded laser returns upon the Earth's surface is computed using a differential global positioning system and an inertial navigation unit. LiDAR systems, with either a small footprint (0.1–0.3 m) or a large footprint (8–70 m), have been used to estimate forest inventory attributes (Lefsky *et al.*, 1999, 2002a; Means *et al.*, 2000; Næsset, 2002; Lim *et al.*, 2003). However, most commercial LiDAR systems for operational purposes are small-footprint, discrete-return airborne lasers, also referred to as airborne laser scanning (ALS). These systems are equipped with a scanning device that distributes the emitted light across a swath width along the aircraft's flight path (Maltamo *et al.*, 2014). As from one to a few (typically four to

five) echoes for each emitted laser pulse can be recorded by ALS, the forest canopy structure can be accurately characterized (Vosselman and Maas, 2010). By distinguishing among the different reflections, the ALS system is capable of penetrating through vegetation and recording the terrain beneath it (Kraus and Pfeifer, 1998). Accordingly, it is possible to generate a Digital Elevation Model (DEM) by filtering ALS data and interpolating the points classified as ground. These DEMs are commonly used to normalize the heights of the ALS point cloud by subtracting the terrain heights to the points (García *et al.*, 2012; Gonzalez-Ferreiro *et al.*, 2012; Montealegre *et al.*, 2014). Thus, this information is potentially more useful for forest inventories than information provided by other remote sensing techniques (Maltamo *et al.*, 2014). Depending on the available point density of the ALS data, two basic inventory techniques can be applied: (1) a tree-based approach with sampling rates larger than 4 points m⁻² (Maltamo *et al.*, 2004; Popescu and Zhao, 2008; Véga and Durrieu, 2011; Li *et al.*, 2012) or (2) an area-based approach (Næsset, 1997b, 2002; Means *et al.*, 2000; Maltamo *et al.*, 2014; Bouvier *et al.*, 2015). With the latter methodology, stand-level attribute estimations are computed from the statistical relationship between plot-level ALS metrics and stand attributes derived from field plots (Bouvier *et al.*, 2015). The pool of commonly used potential predictors derived from ALS data basically belong to two distinct categories (Næsset, 2002; Maltamo *et al.*, 2014; Bouvier *et al.*, 2015):

(1) canopy height-related variables such as mean height and (2) canopy density metrics that characterize the frequency of echoes above a certain height threshold relative to the total number of recorded returns. Metrics that provide the greatest explanation are then selected, with only a few remaining in the final model (Næsset, 2002; Lim et al., 2003; Patenaude et al., 2004; Hall et al., 2005).

This approach has led to a huge number of different models for variables commonly used in forest planning (such as mean height, mean diameter, squared mean diameter, stand basal area, timber volume, stand density or biomass), developed in forests of different characteristics and species (Næsset, 1997a, b, 2002; Magnussen et al., 1999; Means et al., 2000; Næsset and Bjerknes, 2001; Holmgren et al., 2003; Lim et al., 2003; Næsset et al., 2004; Hall et al., 2005; Coops et al., 2007; Næsset and Gobakken, 2008; Zhao et al., 2009; Kronseder et al., 2012; Treitz et al., 2012; Watt et al., 2013; Watt and Watt, 2013; Bouvier et al., 2015). In this regard, correlations of moderate to high strength have been found between ALS metrics and the aforementioned structural variables, except for stand density, which is typically predicted with a moderate to low degree of precision.

Despite the wide acceptance of the use of ALS technology to estimate forest structural variables, little research has focussed on the characterization of Aleppo pine (*Pinus halepensis* Mill.) forest at stand level using this technology, particularly using ALS data with low point density as the one provided by the Spanish National Plan for Aerial Orthophotography (PNOA) with 0.5 points m⁻². In addition, the Mediterranean pine forest differs in structure and species composition from boreal forests, in which this methodology is fully developed (Maltamo et al., 2014). In this sense, the application of ALS technology to low-density Mediterranean pine forests is of scientific interest as trees present irregular crowns, are more dispersed and their leaf area is lower than in boreal forests, which can affect the application of ALS.

Aleppo pine can be found naturally and in plantations, all around the Mediterranean basin, and is well adapted to fire-prone habitats due to its high post-fire seeding capacity (Moya et al., 2008; Chuvieco, 2009; Fournier et al., 2013). According to the Third Spanish National Forest Inventory (NFI 3), pine forest is the most significant habitat in the Autonomous Region of Aragón (north-east Spain) as it represents 49.88 per cent of forested area.

In this regard, the main objective of this study is to test the validity of linear, power function and exponential models in order to estimate stand variables, such as mean height, squared mean diameter, basal area, timber volume, stand density and crown ratio, in a Mediterranean forest of *P. halepensis* Mill, located in Aragón Region, using the discrete, multiple-return, low point density PNOA-ALS data. The estimation of such structural variables is required for designing and planning silvicultural treatments appropriately. A second objective, implicit in the first one, is the selection of ALS-based metrics useful as predictors of the stand-level structural variables of interest.

Materials and methods

Study area

The area of interest is located in the central Ebro valley, which crosses the Autonomous Region of Aragón, in north-eastern Spain (41° 50' N, 0° 57' W) (Figure 1). The Ebro Basin constitutes the northernmost semi-arid region in

Europe and stretches from the Pyrenees range, in the north, to the Iberian range, in the south. Climate of the region is Mediterranean with continental features. Annual precipitation is low and irregular, averaging 350 mm and mostly occurring in autumn and spring. Winters have a monthly mean temperature of ~7°C, and hot, dry summers have temperatures of ~24°C (Vicente-Serrano et al., 2010).

The area is fragmented in stands of variable sizes and occupies 8266 ha. The majority is occupied by natural forest; however, some of the stands, particularly the ones located in the south-southeastern part of the study area, are planted and have an average of 40 years. Most of the study area is located inside the Military Training Center (CENAD) 'San Gregorio', where most stands are owned by the Spanish Ministry of Defense. Accordingly, silvicultural treatments are only focussed on thinning and understorey removal at the roadsides.

In some areas, Aleppo pine forest is interspersed with evergreen shrub species that also compose its understorey. These bushes are dominated by *Quercus coccifera* L., *Juniperus oxycedrus* L. subsp. *macrocarpa* (Sibth. & Sm.) Ball, *Thymus vulgaris* L. and *Rosmarinus officinalis* L. In general, this pine forest shows a regular structure, with a high percentage of codominant trees due to the relatively wide spacing of the Aleppo pines.

The site is characterized by nutrient-poor, gypsiferous soils and a hilly topography, with elevations ranging from 400 to 750 m a.s.l.

Field plot data

Field data were obtained in 45 plots and served as ground truth in order to adjust and validate the predictive models generated from the ALS metrics. First, a set of x and y coordinates were selected within the bounds of the Aleppo pine stands using a stratified random sampling technique to ensure that the 45 locations covered the range of terrain slopes and canopy cover of the study area (Figure 1). Both variables were estimated from ALS data. Then, the field data were collected from July to September 2014. The centre of the circular plots (15 m radius) was located in the field using a Leica VIVA GS15 CS10 GNSS real-time kinematic Global Positioning System. The average accuracy of the planimetric coordinates was 0.15 m. The total tree height (*h*) and the green crown height were measured using a Vertex instrument for precise height measurement (Haglöf Sweden®). Tree diameters were calculated as breast height diameter (dbh) at the standard height of 1.3 m used in Europe, using a Mantax Precision Blue diameter caliper (Haglöf Sweden®). Additionally, qualitative information was collected, such as understorey species and their percentage cover, and tree status. It should be noted that all trees with a dbh > 7.5 cm were inventoried in each plot. A total of 2063 trees were measured, from which was calculated for each plot, the mean height (*Hm*) as the arithmetic mean value of all tree heights; the squared mean diameter (*Dg*, equation 1); the basal area (*G*, equation 2); the timber volume (*V*, equation 3), considering the stem but excluding branches and using the allometric equation developed in the Spanish NFI 3 for Aleppo pine forest located in the province of Zaragoza (Aragón); stand density (*S*, equation 4) and the crown ratio (*CR*, equation 5), i.e. the ratio of live crown length (the total crown height) to tree height:

$$Dg \text{ (cm)} = \sqrt{\frac{4 \times G}{S \times \pi} \times 10000} \quad (1)$$

$$G \text{ (m}^2 \text{ ha}^{-1}) = \frac{\pi \sum dbh_i^2}{4a} \quad (2)$$

$$V \text{ (m}^3 \text{ ha}^{-1}) = \frac{\sum 0.00094 \times dbh_i^{1.9209} \times h_i^{0.7264}}{a} \quad (3)$$

$$S \text{ (stems ha}^{-1}) = \frac{N}{a} \quad (4)$$

$$CR \text{ (\%)} = \frac{L}{Hm} \times 100 \quad (5)$$

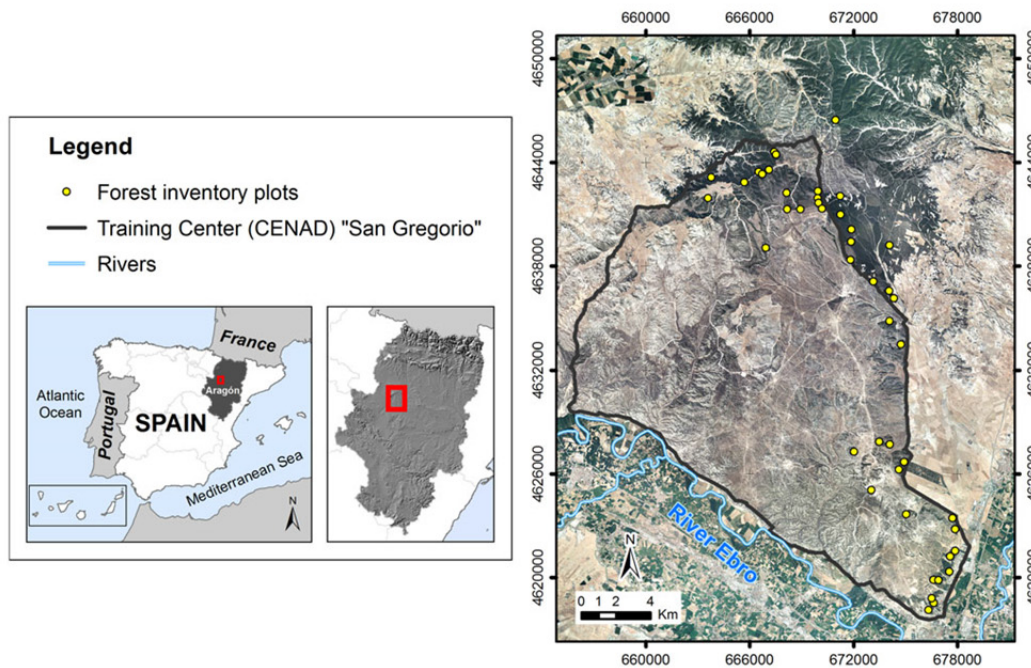


Figure 1 Study area with the locations of the 45 forest inventory plots overlaid over the high spatial resolution orthophotography (PNOA-2009) used as backdrop.

where dbh_i is the diameter at breast height of each tree inventoried expressed in m for G, and in mm for V; a is the plot area (0.070686 ha); h is the tree height in m; N is the total number of trees inside the plot and L is the mean live crown length in m. It should be noted that CR values range from 0 per cent (no crown, dead or defoliated) to 100 per cent (crown extends over the entire tree bole) (Smith, 1997). This ratio is widely used to predict growth and yield of trees and forests, and it is a useful indicator of tree vigour, wood quality, stand density, competition and survival potential, resistance to wind damage, and is a feature of interest in management of wildlife habitat and visual quality (Temesgen *et al.*, 2005).

ALS data

The ALS data were provided by the PNOA (http://www.ign.es/PNOA/vuelo_lidar.html) and captured in several surveys conducted between 23 January and 2 February 2011, using a small-footprint, airborne Leica ALS60 discrete-return sensor. Data were delivered in 2×2 km tiles of raw data points in LAS binary files format v. 1.2, containing x , y and z coordinates (UTM Zone 30 ETRS 1989), with up to four returns measured per pulse. The flying height of the ALS mission was ~ 3000 m a.g.l. The ALS60 sensor was operating in $1.064 \mu\text{m}$ wavelength, 0.22 mrad beam divergence and ± 29 scan angle degrees from nadir. The resulting ALS point density of the study area was 1 point m^{-2} , considering all returns with a vertical accuracy better than 0.20 m.

Ground points were classified using the multiscale curvature classification algorithm (Evans and Hudak, 2007), implemented in the MCC 2.1 command-line tool, and specifically developed to process discrete-return ALS data in forested environments (Montealegre *et al.*, 2015). The Point-TIN-Raster interpolation method (Renslow, 2013), implemented in ArcGIS 10.2 software (ESRI, Redlands, CA, USA), was applied to the ground points to produce a DEM with a 1-m cell grid. The ground elevation

value of the DEM was subtracted from the ALS point height in order to obtain the aboveground heights (i.e. normalized heights) using FUSION LDV 3.30 open source software (McGaughey, 2009). Then, the normalized ALS point clouds were clipped to the spatial extent of each field plot, and a full suite of statistical metrics commonly used as independent variables in vegetation modelling was generated (Evans *et al.*, 2009). They included a range of ALS height percentiles from the 1st (P_{01}) to the 99th (P_{99}); several metrics describing the height distribution of the returns, i.e. elevation minimum and maximum, mean, mode, standard deviation (SD), variance, coefficient of variation (CV), interquartile distance (IQ), skewness and kurtosis; and the percentage of first or all returns above a height threshold. Following Nilsson (1996) and Naesset and Okland (2002), ALS-derived metrics were calculated after applying a threshold value of 1 m height for removing ground and understorey laser hits.

Selection of ALS metrics

A multivariate linear regression approach was adopted to develop models for stand attribute predictions according to Means *et al.* (1999), Naesset (2002), Lim *et al.* (2003), Gonzalez-Ferreiro *et al.* (2012), Garcia *et al.* (2012) and Watt *et al.* (2013). Data processing was performed in the R statistical environment (<http://www.r-project.org/>). Given the large number of potential ALS-derived metrics, the Spearman's rank correlation coefficient (ρ) was first applied in order to select those independent variables with the strongest correlation coefficient with field plot data. A minimum ρ value of ± 0.5 was selected as a threshold to select the ALS metrics used in the next step of the methodology.

Establishment of the models

The selected variables were included in a forward stepwise regression, trying to develop a parsimonious model in order to avoid overfitting

(Hair *et al.*, 1999; Andersen *et al.*, 2005; Chen *et al.*, 2007; González-Olabarria *et al.*, 2012; Maltamo *et al.*, 2014). Predictor variables with a significance value of partial *F* statistic greater than 0.05 were removed from the models (Naesset and Okland, 2002). Furthermore, it was verified whether the fitted model met the statistical assumptions of linearity, normality of the residuals, homoscedasticity and independence or no autocorrelation in the residuals (González-Olabarria *et al.*, 2012).

In the cases where statistical assumptions of linear models could not be fulfilled, such as homoscedasticity, logarithmic transformations of dependent and independent variables were also explored (Means *et al.*, 2000; García *et al.*, 2012). In the case of existence of an exponential relationship between variables, a linear regression can be used by transforming the dependent variable applying a natural logarithm. The linear model generated by applying inverse logarithmic transformation (exponential conversion) will be equivalent to an exponential model:

$$\ln(y) = \ln(a) + bX \Leftrightarrow Y = ae^{bX} \quad (6)$$

Conversely, when all variables are transformed using the natural logarithm, the linear model generated by applying inverse logarithmic transformation will be equivalent to a potential model:

$$\ln(y) = \ln(a) + b \ln(X) \Leftrightarrow Y = aX^b \quad (7)$$

In both cases, the logarithmic transformation introduces a systematic bias in the model, as the additive error generated in the linear model becomes multiplicative applying this transformation (Finney, 1941; Baskerville, 1972; Sprugel, 1983). Thus, according to the protocol used by Hall *et al.* (2005), it is necessary to back-transform the final model applying a correction factor (CF), which is calculated with the following expression (Sprugel, 1983):

$$CF = \exp\left(\frac{SEE^2}{2}\right) \quad (8)$$

where SEE is the standard error of regression estimates and is calculated from the following equation:

$$SEE = \sqrt{\frac{\sum_{i=0}^n (\ln Y_i - \ln \hat{Y}_i)^2}{n - p}} \quad (9)$$

where *p* is the number of parameters in the final model, *n* is the number of field plots, $\ln Y_i$ are the values of the dependent variable and $\ln \hat{Y}_i$ are the corresponding predicted values.

Model validation

A leave-one-out cross-validation (LOOCV) technique was applied to evaluate the accuracy of each predictive model (Andersen *et al.*, 2005; Bouvier *et al.*, 2015). In this method, one of the observations (i.e. a field plot) is removed from the dataset at a time, and the selected model is fitted to the data from the remaining plots. The dependent variable is then predicted for the removed observation. This procedure is repeated until predicted values are obtained for all plots (Naesset and Okland, 2002). Then, the model coefficients from all of the iterations are averaged to obtain the final model. Similarly, LOOCV enables the assessment of model goodness of fit by averaging statistical estimators of model accuracy that were computed at each step (Bouvier *et al.*, 2015).

The predictive value of each regression model was evaluated comparing the adjusted coefficient of determination (R^2), which is the fraction of variance that is explained by the model. In addition, the root-mean-square error (RMSE) was compared against the root-mean-square error for cross-validation (RMSEcv) (Andersen *et al.*, 2005). The RMSE gives an idea of the

Table 1 Summary of the field plot data for structural variables (*n* = 45)

	Min.	Max.	Range	Mean	SD
Hm (m)	3.7	11.3	7.6	7.1	2.0
Dg (cm)	8.7	29.1	20.4	15.8	5.1
G (m ² ha ⁻¹)	0.4	32.7	32.3	13.2	7.4
V (m ³ ha ⁻¹)	0.9	147.7	146.7	49.3	35.4
S (stems ha ⁻¹)	70.7	1273.2	1202.5	671.2	268.1
CR (%)	47.9	95.1	47.2	72.4	13.1

precision estimates in the same units as the dependent variable. A close agreement between RMSEcv and RMSE indicates that the regression model is not overfitting the data and presents good predictive value (Andersen *et al.*, 2005). Models with transformed variables were validated keeping their linear form, without performing the inverse logarithmic transformation described above. Finally, model bias was examined by plotting reference values against predicted values (Watt *et al.*, 2013) and quantitatively calculating the mean of the residuals in each model. It should be noted that in order to achieve comparability, the bias and the RMSE values of models whose variables were logarithmically transformed must be recalculated after applying the inverse logarithmic transformation model as they are not in the same units as the variable to be estimated (García *et al.*, 2012).

Results

A summary of the structural variables estimated in the field is presented in Table 1. As can be observed in the range values, there are a significant variation in all of them, particularly in S.

Table 2 shows the Spearman's rank correlation coefficients between plot-derived metrics (Hm, Dg, G, V, S and CR) and ALS-derived metrics, respectively.

Hm variable was strongly correlated to height percentiles as P₅₀ (0.93) and to vertical distribution metrics, such as Elev. mean (0.92). As can be observed in Figure 2, the aforementioned independent variables are positively related to the Hm of the plot.

Dg was strongly correlated to the upper ALS height percentiles (0.89 in all of them), as well as to Elev. mean (0.88) and Elev. maximum (0.87) metrics (Table 2). As can be seen in Figure 2, the greater the height of the trees within the plot, the greater the Dg value, thus determining the strong correlation with height-related metrics. Furthermore, the positive high correlation with ALS metrics such as Elev. SD (0.79) suggests that the Dg increases with a higher variability of the returns height. However, Elev. skewness presents an inverse correlation (−0.75), indicating a concentration of points above the mean height (Figure 2c). On the contrary, a positive value of Elev. skewness indicates a higher concentration of points of low height values (Figure 2a).

In the case of G, in addition to canopy height-related variables (i.e. P₄₀), the percentage of first returns above mean height presents a high correlation coefficient (0.87) (Table 2). As can be seen in Figure 2, the higher percentage of first returns above 1 m the greater the area occupied by trees stratum.

Following the same tendency, V shows an inverse relationship to Elev. skewness (−0.83). This is observed in the Figure 2a–c. The higher the height of the tree returns, implying a higher V, the lower the Elev. skewness value (negative values). Another important metric for V value is the percentage of first returns above mean

Use of low point density ALS data

Table 2 Correlation coefficients (ρ) describing the strength of linear relationships between plot-derived Hm, Dg, G, V, S and CR, and ALS-derived metrics

Variable	Hm	Dg	G	V	S	CR
Canopy height percentile metrics (m)						
P ₀₁	0.40**	0.20 ^{ns}	0.47**	0.46**	0.56**	-0.33*
P ₀₅	0.65**	0.47**	0.71**	0.71**	0.49**	-0.55**
P ₁₀	0.81**	0.69**	0.87**	0.87**	0.34*	-0.73**
P ₂₀	0.91**	0.81**	0.94**	0.95**	0.20 ^{ns}	-0.78**
P ₂₅	0.92**	0.84**	0.95**	0.96**	0.17 ^{ns}	-0.79**
P ₃₀	0.92**	0.85**	0.95**	0.96**	0.15 ^{ns}	-0.78**
P ₄₀	0.92**	0.87**	0.95**	0.96**	0.12 ^{ns}	-0.78**
P ₅₀	0.93**	0.89**	0.94**	0.96**	0.09 ^{ns}	-0.76**
P ₆₀	0.92**	0.89**	0.93**	0.95**	0.07 ^{ns}	-0.76**
P ₇₀	0.91**	0.89**	0.93**	0.95**	0.06 ^{ns}	-0.76**
P ₇₅	0.91**	0.89**	0.92**	0.95**	0.06 ^{ns}	-0.76**
P ₈₀	0.91**	0.89**	0.92**	0.95**	0.06 ^{ns}	-0.76**
P ₉₀	0.90**	0.88**	0.91**	0.94**	0.05 ^{ns}	-0.77**
P ₉₅	0.88**	0.87**	0.88**	0.90**	0.01 ^{ns}	-0.74**
P ₉₉	0.86**	0.88**	0.86**	0.89**	-0.01 ^{ns}	-0.74**
Canopy height metrics (m)						
Elev. minimum	0.10 ^{ns}	-0.09 ^{ns}	0.13 ^{ns}	0.13 ^{ns}	0.40**	0.02 ^{ns}
Elev. maximum	0.87**	0.87**	0.85**	0.88**	-0.01 ^{ns}	-0.72**
Elev. mean	0.92**	0.88**	0.94**	0.96**	0.10 ^{ns}	-0.77**
Elev. mode	0.87**	0.83**	0.90**	0.92**	0.10 ^{ns}	-0.77**
Canopy height variability metrics (m)						
Elev. SD	0.71**	0.79**	0.67**	0.71**	-0.18 ^{ns}	-0.60**
Elev. variance	0.71**	0.79**	0.67**	0.71**	-0.18 ^{ns}	-0.60**
Elev. CV	-0.31*	-0.11 ^{ns}	-0.43**	-0.40**	-0.51**	0.30*
Elev. IQ	0.56**	0.63**	0.49**	0.53**	-0.22 ^{ns}	-0.39**
Elev. skewness	-0.77**	-0.75**	-0.83**	-0.83**	-0.10 ^{ns}	0.65**
Elev. kurtosis	0.48**	0.49**	0.53**	0.52**	0.06 ^{ns}	-0.53**
Canopy density metrics (%)						
% first returns above 1 m	0.63**	0.55**	0.84**	0.82**	0.44**	-0.79**
% all returns above 1 m	0.45**	0.33*	0.65**	0.63**	0.51**	-0.60**
% first returns above mean height	0.70**	0.65**	0.87**	0.86**	0.34*	-0.80**
% first returns above mode height	0.26 ^{ns}	0.15 ^{ns}	0.42**	0.38**	0.40**	-0.40**
% all returns above mean height	0.61**	0.53**	0.78**	0.77**	0.39**	-0.70**
% all returns above mode height	0.06 ^{ns}	-0.05 ^{ns}	0.19 ^{ns}	0.16 ^{ns}	0.35*	-0.17
(All returns above 1 m)/(total first returns) × 100	0.65**	0.57**	0.85**	0.83**	0.43**	-0.80**
(All returns above mean)/(total first returns) × 100	0.70**	0.65**	0.87**	0.86**	0.34*	-0.80**
(All returns above mode)/(total first returns) × 100	0.26 ^{ns}	0.15 ^{ns}	0.42**	0.38**	0.40**	-0.40**

** $P < 0.01$; * $P < 0.05$; ns, not significant ($P > 0.05$).

height (0.86), which represents the surface occupied by the tree canopy with respect to the ground.

S exhibited weakly significant correlations with height percentiles in comparison with the other variables. An important metric is the percentage of all returns above 1 m (0.51), which denotes the canopy cover.

Finally, CR presents significant negative correlations with ALS metrics, such as P₂₅ (-0.79) and the percentages of returns above 1 m or above mean height. As can be observed in Figure 2a, Aleppo pines with low height values (below 4 m) have the first living branch emerging from the lower part of the stem. However, in taller trees, the first living branch is located in the upper part of the stem (see Figure 2c).

As can be observed in Table 3, Hm variable was best estimated by a simple linear relationship with P₉₉. This relationship accounted for 87 per cent of the variance in the data with a RMSE and RMSEcv of 0.72 m in both cases. The best model for Dg included P₅₀, and Elev. SD with positive linear terms accounted for 84 per cent of the variance in the data. The RMSE and RMSEcv were 1.99 cm. G was best predicted by Elev. maximum and percentage of first returns above 1 m with positive linear relationships, and Elev. skewness with a negative linear relationship. The R^2 was 0.89 and the RMSE and RMSEcv were 2.39 and 2.40 m² ha⁻¹, respectively. The model is coherent since it includes canopy height and variability metrics, which are related to stem diameters. The model selected to estimate V included P₉₅, Elev. kurtosis and percentage of all

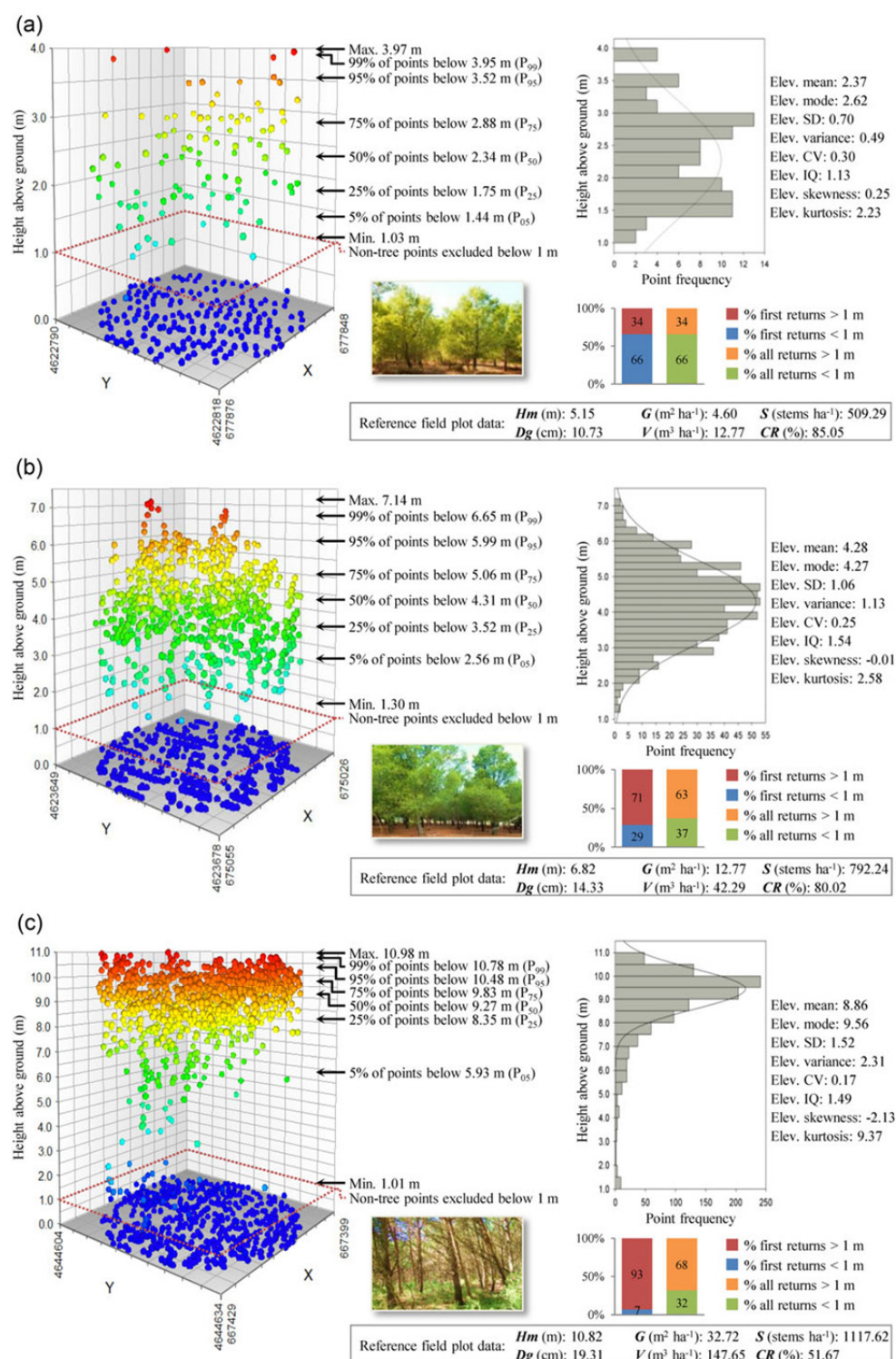


Figure 2 Metrics associated with the vertical distribution of ALS returns in three selected field plots representative of the whole diversity of the pine forest under research: (a) shorter Aleppo pines with open areas, (b) medium height Aleppo pines and (c) taller Aleppo pines with some understorey.

Table 3 Summary of selected models and the validation results for the estimated variables

Dependent variable	Predictive model	Fitting phase			Cross-validation		
		R^2	RMSE	Bias	R^2	RMSEcv	Bias
Hm (m)	$1.881 + 0.751 \times P_{95}$	0.87	0.72	0.00	0.87	0.72	0.00
Dg (cm)	$4.563 + 1.066 \times P_{50} + 4.921 \times \text{Elev. SD}$	0.84	1.99	0.00	0.84	1.99	-0.02
G ($\text{m}^2 \text{ha}^{-1}$)	$-4.370 + 1.134 \times \text{Elev. maximum} - 4.322 \times \text{Elev. skewness} + 0.129 \times \% \text{ first returns above 1 m}$	0.89	2.39	0.00	0.89	2.40	0.03
V ($\text{m}^3 \text{ha}^{-1}$)	$-67.291 + 9.706 \times P_{95} + 8.178 \times \text{Elev. kurtosis} + 0.517 \times \% \text{ all returns above 1 m}$	0.89	11.01	0.00	0.89	10.99	0.03
S (stems ha^{-1})	$496.199 - 1391.986 \times \text{Elev. CV} + 10.383 \times \% \text{ all returns above 1 m}$	0.48	187.53	0.00	0.48	187.32	-0.91
ln CR (%)	$5.576 - 0.169 \times \ln \text{Elev. maximum} - 0.225 \times \ln \text{Elev. kurtosis} - 0.182 \times \ln \% \text{ first returns above 1 m}$	0.69	0.10	0.00	0.60	0.10	0.00

ln denotes that the variable was logarithmic transformed.

returns above 1 m. The strongest contributors to the model were P_{95} and Elev. kurtosis as descriptors of the development of trees and the forest canopy. The percentage of returns above 1 m also presented a positive linear term. In this sense, areas with larger trees and higher volume allow proportionally less penetration to the ground as they have more biomass to intercept the laser pulse. This model presented a R^2 of 0.89 and the RMSE and the RMSEcv were 11.01 and 10.99 $\text{m}^3 \text{ha}^{-1}$, respectively (Table 3). The model with the highest R^2 for S included Elev. CV with an inverse relationship, and percentage of all returns above 1 m, with a positive relationship. Elev. CV characterizes tree height heterogeneity and the percentage of returns above 1 m is a good indicator of the horizontal distribution of the canopy, both related with the stand density. The total amount of variance accounted for by the multiple regression model of S was markedly lower at 0.48. It presented a RMSE of 187.53 and a RMSEcv of 187.32 stems ha^{-1} . Although the accuracy of the model is relatively low and the error is somewhat high, it should be noted that the number of stems per hectare illustrates the great variability in the study area (the SD in the 45 plots was 268.1 stems ha^{-1} according to Table 1). Furthermore, the failure of ALS height metrics to characterize S is very common. With respect to the CR, it was necessary to transform the independent variables in order to stabilize the variance and to achieve linearity for the model residuals. The final model included the Elev. maximum, the Elev. kurtosis and the percentage of first returns above 1 m with negative linear terms, accounting for 69 per cent of the variation in the data. The RMSE and the RMSEcv presented the same value, 0.10 (7.26 per cent in transformed units) (Table 3). It is noteworthy that while Elev. maximum and the Elev. kurtosis give information about the height of the trees, the percentage of first returns above 1 m describes the canopy density. In this sense, a low percentage indicates a tall green crown, whereas a high percentage indicates a compact and short crown (see Figure 2a,c).

Overall, with the exception of S characteristic, the coefficients of determination for the final models were reasonably high, ranging from 0.69 to 0.89 (Table 3). In addition, the errors were relatively low for all variables apart from S. Plots of predicted against reference values for all models (Figure 3), except S, showed little apparent bias. As can be observed in Table 3, values of bias obtained after applying LOOCV to the final models for Hm, Dg, G, V, S and CR were 0.00 m,

-0.02 cm, 0.03 $\text{m}^2 \text{ha}^{-1}$, 0.03 $\text{m}^3 \text{ha}^{-1}$, -0.91 stems ha^{-1} and 0.00 (-0.09 per cent in transformed units), respectively.

Discussion

Stand-level inventory is the most widely used method for evaluating current characteristics and providing comprehensive information about the state and dynamics of forest. However, over the last two decades, traditional field-based assessment has been progressively replaced or supplemented with ALS in different types of forests due to the capacity of this technology to capture three-dimensional data over large areas in a timely and economic fashion (Vosselman and Maas, 2010; Maltamo et al., 2014).

This study examined the applicability of low point density PNOA-ALS data to an area-based approach to estimate Hm, Dg, G, V, S and CR in Aleppo pine forest. This type of forest is very prominent in the Spanish Mediterranean landscape. However, there is still little available information of the kind estimated in this study.

The results obtained indicate that PNOA-ALS data can be used to make accurate and robust estimates of several important forest stand variables, except for S. ALS variables selected in the present study to estimate structural variables were similar to those used by other studies, such as Næsset (2002), Li et al. (2008) and García et al. (2012), which followed an analogous approach. In general, the adjusted models included at least one ALS-derived metric related to the canopy height distribution and another concerning the horizontal distribution of the point cloud, i.e. the canopy density or the percentage of returns above a certain height threshold. It should be noted that previous studies usually performed a logarithmic transformation of variables. However, our results suggest that it is not necessarily the best option. In fact, it was only required to fulfil the assumptions of linear regression model in the case of CR estimation.

Of the structural variables commonly estimated in forest inventory, *a priori* Hm seems to be the one that should be predicted with the highest precision using ALS data. However, the coefficient of determination for the final model was 0.87, slightly lower than the one for V and G variables. Nevertheless, this value falls within the typical values reported by Means et al. (2000), Næsset (2002), Coops et al. (2007), Stone et al. (2011), Gonzalez-Ferreiro et al. (2012), Treitz et al. (2012), Watt et al. (2013) and Watt and

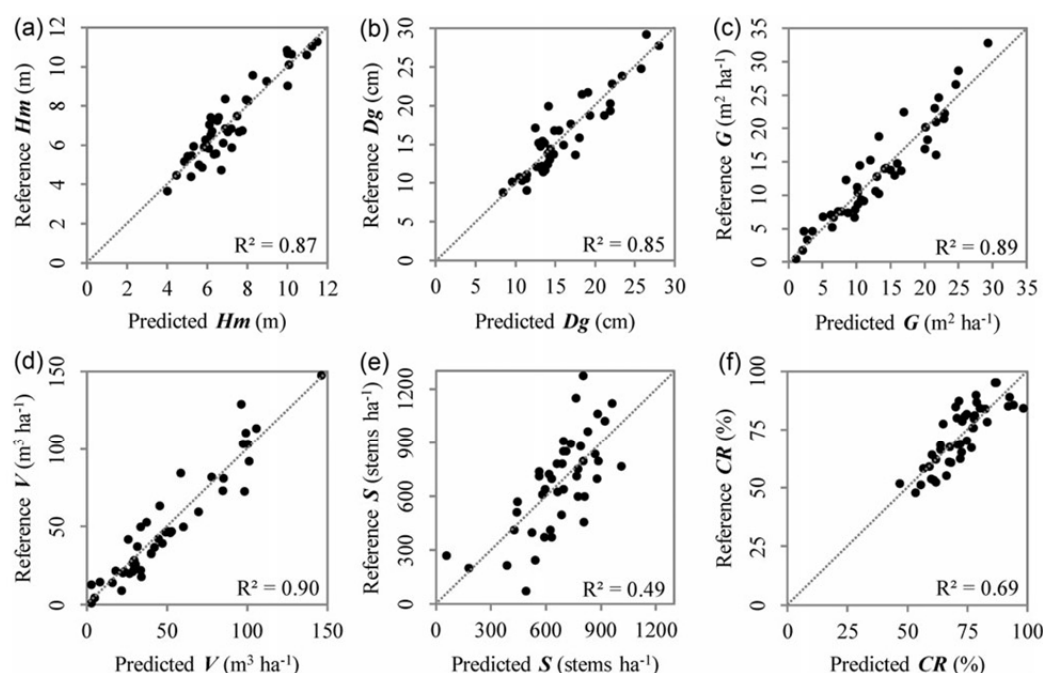


Figure 3 Relationship between predicted and reference (a) Hm, (b) Dg, (c) G, (d) V, (e) S and (f) CR values.

Watt (2013), which ranged from 0.82 to 0.98. It should be noted that the accuracy of the model can depend on the type of coniferous species studied, among other factors. In the case of pine species with a wider crown, such as Aleppo pines, the upper part of the tree is more likely to be recorded by ALS technology (Watt et al., 2013; Maltamo et al., 2014). However, the use of low-density ALS data may lead to an underestimation of tree height (Vosselman and Maas, 2010). Thus, it would be desirable for future PNOA missions to increase point density in order to allow accurate tree top sampling by detecting the uppermost portion of the forest canopy.

V was predicted with a high degree of precision (R^2 of 0.89), coinciding with the range of cited previous studies for coniferous species (from 0.46 to 0.97) (Næsset, 1997b; Hollaus et al., 2007; Watt and Watt, 2013). In this respect, in addition to the errors introduced by the prediction model, those generated in the estimation using the allometric equation for the volume estimation should be considered.

Model prediction for G had a comparable precision (R^2 of 0.89) to other studies within coniferous forests in temperate and boreal regions, whose adjusted R^2 typically ranged from 0.62 to 0.95 (Gonzalez-Ferreiro et al., 2012; Watt et al., 2013).

Little previous research has developed models of either stem diameter or green crown height. In this study, both stand metrics were found to be strongly correlated to ALS metrics (R^2 of 0.84 and 0.69, respectively). The result for Dg is consistent with research on coniferous species within Norway, where coefficients of determination ranged from 0.39 to 0.78 (Næsset, 2002), and with the Douglas-fir plantations in New Zealand, where the R^2 was 0.86 (Watt et al., 2013). The degree of precision of the final model of CR (R^2 of 0.69) exceeds that of models previously developed by Næsset and Okland (2002) for stands dominated by *Picea abies* (L.) Karst. and

Pinus sylvestris L. where the coefficient of determination ranged from 0.47 to 0.60.

S was predicted with a low degree of precision (R^2 of 0.48). Stand density assessment based on height measurements is marred by the uncertainty that is inherent in the relationship between height and stems (Maltamo et al., 2014). In this sense, the use of individual tree delineation may provide a more accurate method for deriving a stand density characterization. Recent research has shown that high-resolution ALS-derived imagery can be useful to accurately isolate trees through specific algorithms (Vosselman and Maas, 2010; Li et al., 2012; Maltamo et al., 2014).

Although the number of plots seems not to be high, it is important to consider that fieldwork is a time-consuming task, leading to an increase of the costs of the study. In this regard, the LOOCV technique was adopted in the validation process in order not to reduce even more the sample.

The time delay, i.e. the 3 years between the ALS data acquisition and the field data collection, was not considered a significant source of error, as the pine forest under study did not change considerably during that period. As mentioned above, thinning operations and understorey removal were only carried out at the roadsides in order to prevent the fire risk.

Bearing in mind that point density has a direct effect on ALS data acquisition costs (Maltamo et al., 2014) and in the reliability of the derived models, our results using low-density ALS data support those presented by Gonzalez-Ferreiro et al. (2012). They pointed out that for forest stand variable estimation, laser pulse density could be reduced to low densities (up to 0.5 points m^{-2}) without significant loss of information for the mean and dominant height, stand basal area, stand volume and stand biomass fractions. Thus, according to our results, the PNOA-ALS data, recently delivered by the Spanish National Geographic Institute, become

an excellent source of information for forest management, especially reducing forest inventory costs. In this regard, our study sheds light on the future possibility of developing a NFI using PNOA-ALS data, as Norway and Finland did. In fact, the low-density ALS data that commonly are used in operational forest inventory usually consist of less than one transmitted laser pulse per square metre (Maltamo *et al.*, 2014).

Considering that the results obtained are strongly site dependent and they could be influenced by the structural complexity of the site, our study provides the necessary motivation to perform future studies along different environments. It would be desirable to estimate forest stand variables from PNOA-ALS data in different Mediterranean species in order to compare prediction accuracies and to increase the body of research in this environment. In this sense, further research should focus on sensitivity analysis of sources of error that affect the accuracy of the predictive models.

Conclusion

This study demonstrates that management-relevant forest stand variables can be modelled with reasonable precision in Mediterranean Aleppo pine forest using low-density laser data. The final models constructed for mean height, squared mean diameter, basal area, timber volume, stand density and crown ratio presented R^2 values of 0.87, 0.84, 0.89, 0.89, 0.48 and 0.69, respectively, with RMSEs of 0.72 m, 1.99 cm, 2.39 m² ha⁻¹, 11.01 m³ ha⁻¹, 187.53 stems ha⁻¹ and 7.26 per cent, respectively. With the exception of stand density, all relationships were relatively unbiased and similar to those reported in the international literature. In most cases, the models include variables belonging to the three different types of ALS measurements: variables related to canopy heights, variables depicting vegetation height variability or return distribution and measurements related to canopy density.

Acknowledgements

The authors would also like to thank the anonymous reviewers for adding quality to the previous versions of this paper. The ALS data were provided by the Spatial Information Centre of Aragón. The authors are grateful to the Training Center (CENAD) 'San Gregorio' for its assistance in the field and for the invaluable technical support.

Conflict of interest statement

None declared.

Funding

This work has been financed by the Government of Aragón, Department of Science, Technology and University (FPI Grant BOA 30, 11/02/2011) and supported by the Research Project of Centro Universitario de la Defensa de Zaragoza (Project No: 2013-04).

References

- Andersen, H.-E., McGaughey, R.J. and Reutebuch, S.E. 2005 Estimating forest canopy fuel parameters using LIDAR data. *Remote Sens. Environ.* **94**, 441–449.

- Baskerville, G.L. 1972 Use of logarithmic regression in the estimation of plant biomass. *Can. J. For. Res.* **2**, 49–53.
- Bouvier, M., Durrieu, S., Fournier, R.A. and Renaud, J.-P. 2015 Generalizing predictive models of forest inventory attributes using an area-based approach with airborne LiDAR data. *Remote Sens. Environ.* **156**, 322–334.
- Chen, Q., Gong, P., Baldocchi, D. and Tian, Y.Q. 2007 Estimating basal area and stem volume for individual trees from LiDAR data. *Photogramm. Eng. Remote Sensing*, **73**, 1355–1365.
- Chuvieco, E. 2009 *Earth Observation of Wildland Fires in Mediterranean Ecosystems*. Springer.
- Coops, N., Hilker, T., Wulder, M., St-Onge, B., Newnham, G., Siggins, A. *et al.* 2007 Estimating canopy structure of Douglas-fir forest stands from discrete-return LiDAR. *Trees Struct. Funct.* **21**, 295–310.
- Evans, J.S. and Hudak, A.T. 2007 A multiscale curvature algorithm for classifying discrete return LiDAR in forested environments. *IEEE Trans. Geosci. Remote Sens.* **45**, 1029–1038.
- Evans, J., Hudak, A., Faux, R. and Smith, A.M. 2009 Discrete return lidar in natural resources: recommendations for project planning, data processing, and deliverables. *Remote Sens.* **1**, 776–794.
- Finney, D.J. 1941 On the distribution of a variate whose logarithm is normally distributed. *Suppl. J. J. R. Stat. Soc.* **7**, 155–161.
- Fournier, T.P., Battipaglia, G., Brosier, B. and Carcaillet, C. 2013 Fire-scars and polymodal age-structure provide evidence of fire-events in an Aleppo pine population in southern France. *Dendrochronologia* **31**, 159–164.
- García, D., Godino, M. and Mauro, F. 2012 *Lidar: Aplicación Práctica Al Inventario Forestal*. Editorial Academia Española, p. 196.
- Gonzalez-Ferreiro, E., Dieguez-Aranda, U. and Miranda, D. 2012 Estimation of stand variables in *Pinus radiata* D. Don plantations using different LiDAR pulse densities. *Forestry* **85**, 281–292.
- González-Olabarria, J.-R., Rodríguez, F., Fernández-Landa, A. and Mola-Yudego, B. 2012 Mapping fire risk in the Model Forest of Urbión (Spain) based on airborne LiDAR measurements. *For. Ecol. Manage.* **282**, 149–156.
- Hair, J.F., Anderson, R.E., Tatham, R.L. and Black, W.C. 1999 *Análisis multivariante*. 5ª edn. Prentice Hall Iberia, p. 832.
- Hall, S.A., Burke, I.C., Box, D.O., Kaufmann, M.R. and Stoker, J.M. 2005 Estimating stand structure using discrete-return lidar: an example from low density, fire prone ponderosa pine forests. *For. Ecol. Manage.* **208**, 189–209.
- Hollaus, M., Wagner, W., Maier, B. and Schadauer, K. 2007 Airborne laser scanning of forest stem volume in a mountainous environment. *Sensors* **7**, 1559–1577.
- Holmgren, J., Nilsson, M. and Olsson, H. 2003 Estimation of tree height and stem volume on plots using airborne laser scanning. *For. Sci.* **49**, 419–428.
- Kraus, K. and Pfeifer, N. 1998 Determination of terrain models in wooded areas with airborne laser scanner data. *ISPRS J. Photogramm. Remote Sens.* **53**, 193–203.
- Kronstedt, K., Ballhorn, U., Böhm, V. and Siegert, F. 2012 Above ground biomass estimation across forest types at different degradation levels in Central Kalimantan Using LiDAR data. *Int. J. Appl. Earth Obs. Geoinform.* **18**, 37–48.
- Leboeuf, A., Fournier, R.A., Luther, J.E., Beaudoin, A. and Guindon, L. 2012 Forest attribute estimation of northeastern Canadian forests using QuickBird imagery and a shadow fraction method. *For. Ecol. Manage.* **266**, 66–74.
- Lefsky, M.A., Cohen, W.B., Acker, S.A., Parker, G.G., Spies, T.A. and Harding, D. 1999 Lidar remote sensing of the canopy structure and biophysical properties of Douglas-fir western hemlock forests. *Remote Sens. Environ.* **70**, 339–361.

- Lefsky, M.A., Cohen, W.B., Harding, D.J., Parker, G.G., Acker, S.A. and Gower, S.T. 2002a Lidar remote sensing of above-ground biomass in three biomes. *Global Ecol. Biogeogr.* **11**, 393–399.
- Lefsky, M.A., Cohen, W.B., Parker, G.G. and Harding, D.J. 2002b Lidar remote sensing for ecosystem studies. *BioScience* **52**, 19–30.
- Le Toan, T., Beaudoin, A., Riou, J. and Guyon, D. 1992 Relating forest biomass to SAR data. *IEEE Trans. Geosci. Remote Sens.* **30**, 403–411.
- Li, Y., Andersen, H.-E. and McGaughey, R. 2008 A comparison of statistical methods for estimating forest biomass from light detection and ranging data. *West. J. Appl. Forest.* **23**, 223–231.
- Li, W., Guo, Q., Jakubowski, M.K. and Kelly, M. 2012 A new method for segmenting individual trees from the lidar point cloud. *Photogramm. Eng. Remote Sens.* **78**, 75–84.
- Lim, K., Treitz, P., Wulder, M., St-Onge, B. and Flood, M. 2003 LiDAR remote sensing of forest structure. *Prog. Phys. Geogr.* **27**, 88–106.
- Magnussen, S., Eggermont, P. and LaRiccia, V.N. 1999 Recovering tree heights from airborne laser scanner data. *Forest Sci.* **45**, 407–422.
- Maltamo, M., Eerikainen, K., Pitkänen, J., Hyyppä, J. and Vehmas, M. 2004 Estimation of timber volume and stem density based on scanning laser altimetry and expected tree size distribution functions. *Remote Sens. Environ.* **90**, 319–330.
- Maltamo, M., Næsset, E. and Vauhkonen, J. 2014 *Forestry Applications of Airborne Laser Scanning: Concepts and Case Studies*. Managing Forest Ecosystems. Springer London, Limited.
- McGaughey, R. 2009 *FUSION/LDV: Software for LIDAR Data Analysis and Visualization*. US Department of Agriculture, Forest Service, Pacific Northwest Research Station.
- Means, J., Acker, S., Harding, D., Blair, J., Lefsky, M., Cohen, W. et al. 1999 Use of large-footprint scanning airborne LiDAR to estimate forest stand characteristics in the Western Cascades of Oregon. *Remote Sens. Environ.* **67**, 298–308.
- Means, J.E., Acker, S.A., Fitt, B.J., Renslow, M., Emerson, L. and Hendrix, C.J. 2000 Predicting forest stand characteristics with airborne Lidar. *Photogramm. Eng. Remote Sens.* **66**, 1367–1371.
- Montealegre, A., Lamelas, M., Tanase, M. and de la Riva, J. 2014 Forest fire severity assessment using ALS data in a Mediterranean environment. *Remote Sens.* **6**, 4240–4265.
- Montealegre, A., Lamelas, M. and de la Riva, J. 2015 A comparison of open source LiDAR filtering algorithms in a Mediterranean forest environment. *IEEE J. Select. Top. Appl. Earth Observ. Remote Sens.* **8**, 4072–4085.
- Moya, D., De las Heras, J., López Serrano, F.R. and Leone, V. 2008 Optimal intensity and age of management in young Aleppo pine stands for post-fire resilience. *For. Ecol. Manage.* **255**, 3270–3280.
- Næsset, E. 1997a Determination of mean tree height of forest stands using airborne laser scanner data. *ISPRS J. Photogramm. Remote Sens.* **52**, 49–56.
- Næsset, E. 1997b Estimating timber volume of forest stands using airborne laser scanner data. *Remote Sens. Environ.* **61**, 246–253.
- Næsset, E. 2002 Predicting forest stand characteristics with airborne scanning laser using a practical two-stage procedure and Field data. *Remote Sens. Environ.* **80**, 88–99.
- Næsset, E. and Bjerknes, K.-O. 2001 Estimating tree heights and number of stems in young forest stands using airborne laser scanner data. *Remote Sens. Environ.* **78**, 328–340.
- Næsset, E. and Gobakken, T. 2008 Estimation of above- and below-ground biomass across regions of the boreal forest zone using airborne laser. *Remote Sens. Environ.* **112**, 3079–3090.
- Næsset, E. and Okland, T. 2002 Estimating tree height and tree crown properties using airborne scanning laser in a boreal nature reserve. *Remote Sens. Environ.* **79**, 105–115.
- Næsset, E., Gobakken, T., Holmgren, J., Hyyppä, H., Hyyppä, J., Maltamo, M. et al. 2004 Laser scanning of forest resources: the nordic experience. *Scand. J. Forest Res.* **19**, 482–499.
- Nilsson, M. 1996 Estimation of tree heights and stand volume using an airborne lidar system. *Remote Sens. Environ.* **56**, 1–7.
- Patenaude, G., Hill, R.A., Milne, R., Gaveau, D.L.A., Briggs, B.B.J. and Dawson, T.P. 2004 Quantifying forest above ground carbon content using LiDAR remote sensing. *Remote Sens. Environ.* **93**, 368–380.
- Popescu, S.C. and Zhao, K. 2008 A voxel-based lidar method for estimating crown base height for deciduous and pine trees. *Remote Sens. Environ.* **112**, 767–781.
- Renslow, M. 2013 *Manual of Airborne Topographic Lidar*. The American Society for Photogrammetry and Remote Sensing.
- Smith, D.M. 1997 *The Practice of Silviculture: Applied Forest Ecology*. Wiley.
- Sprugel, D.G. 1983 Correcting for bias in log-transformed allometric equations. *Ecology* **64**, 209–210.
- Stone, C., Penman, T. and Turner, R. 2011 Determining an optimal model for processing lidar data at the plot level: results for a *Pinus radiata* plantation in New South Wales, Australia. *N. Z. J. Forest. Sci.* **41**, 191–205.
- Tanase, M., de la Riva, J., Santoro, M., Pérez-Cabello, F. and Kasischke, E. 2011 Sensitivity of SAR data to post-fire forest regrowth in Mediterranean and boreal forests. *Remote Sens. Environ.* **115**, 2075–2085.
- Temesgen, H., LeMay, V. and Mitchell, S.J. 2005 Tree crown ratio models for multi-species and multi-layered stands of southeastern British Columbia. *Forest. Chron.* **81**, 133–141.
- Treitz, P., Lim, K., Woods, M., Pitt, D., Nesbitt, D. and Etheridge, D. 2012 LiDAR sampling density for forest resource inventories in Ontario, Canada. *Remote Sens.* **4**, 830–848.
- Véga, C. and Durrieu, S. 2011 Multi-level filtering segmentation to measure individual tree parameters based on Lidar data: application to a mountainous forest with heterogeneous stands. *Int. J. Appl. Earth Observ. Geoinform.* **13**, 646–656.
- Vicente-Serrano, S.M., Lasanta, T. and Gracia, C. 2010 Aridification determines changes in forest growth in *Pinus halepensis* forests under semiarid Mediterranean climate conditions. *Agr. Forest Meteorol.* **150**, 614–628.
- Vosselman, G. and Maas, H.-G. 2010 *Airborne and Terrestrial Laser Scanning*. Whittles Publishing, p. 320.
- Watt, P. and Watt, M. 2013 Development of a national model of *Pinus radiata* stand volume from LiDAR metrics for New Zealand. *Int. J. Remote Sens.* **34**, 5892–5904.
- Watt, M., Meredith, A., Watt, P. and Gunn, A. 2013 Use of LiDAR to estimate stand characteristics for thinning operations in young Douglas-fir plantations. *N. Z. J. Forest. Sci.* **43**, 18.
- Zhao, K., Popescu, S. and Nelson, R. 2009 Lidar remote sensing of forest biomass: a scale-invariant estimation approach using airborne lasers. *Remote Sens. Environ.* **113**, 182–196.

3.4 Estimación del carbono contenido en la biomasa arbórea

RESUMEN

La estimación de la biomasa aérea total (BT) de los árboles es esencial para la cuantificar el carbono almacenado y analizar su utilidad como bioenergía, así como para gestionar los bosques de manera sostenible. Sin embargo, se ha realizado poca investigación centrándose en el uso de la tecnología de escaneo láser aerotransportado en bosques mediterráneos de pino carrasco.

El interés de este trabajo radica en evaluar la idoneidad de los datos de LiDAR del PNOA para estimar y cartografiar la BT y su contenido de carbono en masas de *P. halepensis* Mill. en el entorno del Centro de Adiestramiento “San Gregorio” (Zaragoza).

La BT ha sido calculada a partir de 45 parcelas de campo utilizando ecuaciones alométricas y se ha relacionado, mediante un análisis multivariante de regresión lineal, con una colección de variables independientes extraídas de los datos LiDAR. El modelo predictivo para la BT ha sido validado mediante la técnica LOOCV.

Los resultados demuestran que, a pesar de la baja densidad de puntos de los datos LiDAR-PNOA, el modelo final es suficientemente preciso para ser utilizado en la gestión forestal. Su implementación en un SIG, utilizando una cuadrícula regular con un tamaño de celda de 25 × 25 m para calcular la BT a nivel de masa, ha posibilitado su conversión en contenido de carbono. La altura máxima, la curtosis y el porcentaje de los retornos láser superiores a 1 m fueron las variables independientes LiDAR incluidas en el modelo ajustado, que presenta un R^2 de 0,89.



Assessment of biomass and carbon content in a Mediterranean Aleppo pine forest using ALS data

Antonio Luis Montealegre ^{1,*}, María Teresa Lamelas ^{1,2,†}, Juan de la Riva ^{1,†}, Alberto García-Martín ^{1,2,†} and Francisco Escribano ²

¹ GEOFOREST-IUCA research group, Department of Geography, University of Zaragoza, Pedro Cerbuna 12, 50009, Zaragoza, Spain; E-Mail: tlamelas@unizar.es (M.T.L.); delariva@unizar.es (J.R.); algarcia@unizar.es (A.G.)

² Centro Universitario de la Defensa de Zaragoza, Academia General Militar, Ctra. de Huesca s/n, 50090, Zaragoza, Spain; E-Mail: fescrib@unizar.es (F.E.)

[†] These authors contributed equally to this work.

* Author to whom correspondence should be addressed; E-Mail: monteale@unizar.es; Tel.: +34-876-554-058; Fax: +34-976-761-506.

Published: 22 June 2015

Abstract: Tree biomass estimate is essential for carbon accounting, bioenergy feasibility studies and forest sustainable management. However, little research has been conducted focusing on the use of Airborne Laser Scanning (ALS) technology in Mediterranean Aleppo pine forest. Thus, the availability of ALS information provided by the Spanish National Plan for Aerial Orthophotography (PNOA) determined the main objective of this research. It is our aim to test the suitability of the low point density, discrete, multiple-return, ALS data, to estimate and map the total biomass (TB) and its carbon content in *Pinus halepensis* Mill. forest. Accordingly, TB was calculated in 45 field plots, located north-eastern Spain, using allometric equations. This information was related through a multivariate linear regression analysis with a collection of independent variables extracted from the ALS data. The predictive model was validated using a leave-one-out cross-validation (LOOCV) technique. Then, a regular grid with cell size 25 m x 25 m was generated in order to compute TB at stand level. Afterwards, biomass was transformed to carbon content by using a conversion factor. Maximum height, kurtosis and percentage of returns above 1 m were the ALS metrics included in the fitted model, which presented a R^2 value of 0.89.

Keywords: Airborne laser scanning; LiDAR; total biomass; carbon stock; forest inventory; multivariate linear regression; Mediterranean forest

1. Introduction

In the last decades, operational collection of information on relevant characteristics of forest ecosystems by means of pure ground-based field inventories has been revolutionized by the development of remote sensing sensors [1]. Optical and radar remote sensing have been widely used to map forest structural attributes and biophysical parameters [2-5]. In this sense, Light Detection and Ranging (LiDAR) has emerged as a promising techniques in forest attribute estimation [1,6,7]. Most commercial LiDAR systems are small-footprint, discrete-return airborne lasers, also referred to as Airborne Laser Scanning (ALS). These systems are able to accurately characterize the three-dimensional structure of the forest canopy due to its capacity to record the different reflections of each emitted pulse, corresponding to both, vegetation and terrain beneath it [6,8]. Accordingly, in order to generate a precise Digital Elevation Model (DEM), it is necessary to filter the ALS data and to interpolate the points classified as ground. These DEMs are needed to normalize the heights of the ALS point cloud with respect to the terrain heights [9-11].

Nowadays, the most common methodology in forestry is the area-based one [12-14,1]. In this approach, attributes such as mean height, mean diameter, squared mean diameter, stand basal area, timber volume, stand density or biomass [15-24] are computed at stand level and statistical related to a pool of potential ALS-based predictor variables [14]. In this regard, correlations of moderate to high strength have been found between ALS metrics and the aforementioned dasometric attributes [1].

Particularly, forest biomass mapping has gained interest in recent years due to the following three reasons: i) the relevant role of forests in the carbon cycle and in the balance of greenhouse gases emissions, ii) the assessment of the available resources for bioenergy production and iii) the sustainable management of forests [25]. The total biomass (TB) refers to the dry weight of the plant material from trees, including roots, stems, bark, branches and leaves from the ground to the apex [1]. Conventional methods used to estimate this variable are based on field measurements. However, these methods are limited by the high cost of such destructive sampling, necessary to create allometric models required to extrapolate the results. Alternatively, remote sensing has allowed forest biomass mapping in a wide range of spatial and temporal scales, reducing costs and fieldwork. Numerous studies have correlated the biomass and the spectral response of vegetation using passive optical sensors. However, the results have been typically affected by saturation problems when the biomass is greater than 100 tons/ha [25]. On the other hand, approaches using active sensors such as Synthetic Aperture Radar (SAR) have shown that they are more sensitive to higher levels of biomass and, more recently, the use of ALS technology is becoming a tool with great potential to estimate biomass [26,1]. This fact, together with the availability of ALS data captured for the entire Spanish territory under the National Plan for Aerial Orthophotography (PNOA), has determined the aim of this work, which is to evaluate the suitability of this information to estimate and map TB and carbon content in *Pinus halepensis* Mill. forest at stand level. Our objective implies developing a statistical model that

adequately relates the information provided by the ALS data with the one extracted directly from the field.

2. Methods

2.1. Study area

The study area corresponds to monospecific stands of *Pinus halepensis* Mill. and it is located in the central Ebro valley (41° 50' N, 0° 57' W) (Fig. 2). This river crosses the Autonomous Region of Aragón, sited northeastern Spain. The forest under study is fragmented in stands of variable sizes and occupies 8266 ha. In some areas, Aleppo pine forest is interspersed with evergreen shrubs, dominated by *Quercus coccifera* L., *Juniperus oxycedrus* L. subsp. *macrocarpa* (Sibth. & Sm.) Ball, and *Thymus vulgaris* L. Part of the study area is located inside the Military Training Center (CENAD) “San Gregorio”, involving a direct risk of fire. According to the Third Spanish National Forest Inventory (NFI 3), pine forest is the most significant habitat in Autonomous Region of Aragón. In fact, it represents 49.88% of the forested area. In addition, Aleppo pines play an important role in the protection and restoration of forest as this tree species is practically the only one adapted to the adverse climatic and edaphic conditions of the study area. In this regard, it is important to notice that the study area is characterized by nutrient-poor, gypsiferous soils. In addition, this area presents a hilly topography, with elevations ranging from about 400 m to 750 m above sea level. Climate of the region is Mediterranean with continental features, i.e., irregular annual precipitation, cold winters and hot, dry summers [27].

2.2. ALS data

The ALS data were provided by the Spanish National Plan for Aerial Orthophotography (PNOA, http://www.ign.es/PNOA/vuelo_lidar.html). The information was captured in several surveys conducted between January and February 2011. A Leica ALS60 sensor was used. This is a small-footprint, discrete return sensor. Data were delivered in 2 km x 2 km tiles of unclassified points in LAS binary file, format v. 1.2, containing *x*, *y*, *z* coordinates (UTM Zone 30 ETRS 1989), with up to four returns recorded per pulse. The flying height of the ALS mission was around 3,000 m.a.g.l. The sensor was operating in 1.064 μ m wavelength, 0.22 mrad beam divergence and ± 29 scan angle degrees. The resulting ALS point density for the study area was 1 point/m² with a vertical accuracy higher than 0.20 m.

After eliminating the noise of the point cloud, it was necessary to perform a filtering process to identify ground returns. The classification algorithm implemented in the MCC software 2.1, developed by Evans and Hudak [28], was applied. This algorithm was selected given its good performance in hilly forest landscapes [9]. Points classified as ground were interpolated using the “Point-TIN-Raster” interpolation method [29] in ArcGIS 10.2 software (ESRI, Redlands, CA, USA) to create a DEM. This model, with 1 m spatial resolution, was used to normalize the heights of the point cloud. Then, using the “ClipData” and “CloudMetrics” commands implemented in FUSION LDV 3.30 software [30], several ALS point cloud metrics commonly used in vegetation structure modeling were calculated at each plot. A threshold of 1 m above the ground was applied in the selection of the points used in the

calculation in order to exclude from the metrics the returns belonging to bare-earth and understory [31].

2.3. Field data

The fieldwork campaign was conducted between July and September 2014. The location of 45 circular plots with 15 m radius was selected via a stratified random sampling, trying to obtain a representative sample of the variability of forest height and topographic complexity [31]. The selected sites were staked out in field using a Leica VIVA GS15 CS10 real-time kinematic Global Navigation Satellite System, achieving an average planimetric accuracy of 0.15 m. The total tree height (h) was measured using a *Vertex* instrument for precision height (Haglöf Sweden®). In addition, the tree diameter at breast height (dbh), considering the standard height of 1.3 m used in Europe, was measured using a calliper *Haglöf Mantax Precision Blue* (Haglöf Sweden®). In total 2,063 trees were inventoried (only trees greater than 7.5 cm dbh). The different biomass fractions at each plot was calculated using the *Pinus halepensis* Mill allometric models reported by Ruiz-Peinado *et al.* [32] (equations 1, 2, 3, 4 and 5):

$$W_s = 0.0139 \cdot dbh^2 \cdot h \quad (1)$$

$$W_{b7} = [3.926 \cdot (dbh - 27.5)] \cdot Z; \text{ If } dbh \leq 27.5 \text{ cm then } Z = 0; \text{ If } dbh > 27.5 \text{ cm then } Z = 1 \quad (2)$$

$$W_{b2-7} = 4.257 + 0.00506 \cdot dbh^2 \cdot h - 0.0722 \cdot dbh \cdot h \quad (3)$$

$$W_{b2+n} = 6.197 + 0.00932 \cdot dbh^2 \cdot h - 0.0686 \cdot dbh \cdot h \quad (4)$$

$$W_r = 0.0785 \cdot dbh^2 \quad (5)$$

Where W_s is the biomass weight of the stem fraction (kg), W_{b7} is the biomass weight of the thick branch fraction (diameter larger than 7 cm) (kg), W_{b2-7} is the biomass weight of medium branch fraction (diameter between 2 and 7 cm) (kg), W_{b2+n} is the biomass weight of thin branch fraction (diameter smaller than 2 cm) with needles (kg), and W_r is the biomass weight of the belowground fraction (kg).

The results of these equations were aggregated to calculate the TB in kilograms of each tree in the field plot, in order to extrapolate the TB values from sample plots to per hectare biomass value (kg of dry biomass per ha).

2.4. Predictive model for estimating the TB

Following Means *et al.* [12], Næsset [13], Gonzalez-Ferreiro *et al.* [11] and García *et al.* [10], a multivariate linear regression approach was adopted in order to develop a model able to estimate the TB. The statistical analysis was performed in the *R* environment (<http://www.r-project.org/>). Given the large number of potential ALS-derived metrics, the Spearman's rank correlation coefficient (Rho) was previously applied in order to select those independent variables with the strongest correlation coefficient with TB [22]. Then, the selected variables were included in a stepwise regression, trying to develop a parsimonious model in order to avoid over-fitting [33,34]. Predictor variables with a

significance value of partial F statistic greater than 0.05 were removed from the model [31]. Besides, the fitted model was selected in compliance with the statistical assumptions of linearity, normality of the residuals, homoscedasticity, and independence or no autocorrelation [10].

A leave-one-out cross-validation (LOOCV) technique was applied with the purpose of performing an unbiased assessment of the predictive capacity of the model [14]. The model goodness-of-fit was assessed comparing the adjusted coefficient of determination (R^2), the root-mean-square error (RMSE) and the mean of the residuals obtained from LOOCV technique [35]. The final model was obtained by averaging the model coefficients computed at each step of the cross-validation performed [14].

2.5. Mapping TB and carbon content

The pixel size selected to compute the ALS-derived metrics and to map TB was 25 m x 25 m, representing an area of 625 m², similar to the field plot dimensions (706.86 m²). Metrics involved in the equation of the TB model were obtained using the “GridMetrics” and “CSV2Grid” commands implemented in FUSION LDV 3.30 [30]. Then, using map algebra in ArcGIS 10.2 (ESRI, Redlands, CA, USA), model coefficients were applied to the raster layers to achieve a TB map at stand level. Finally, TB values were converted to carbon content using a conversion factor of 0.499 following Montero *et al.* [36].

3. Results and Discussion

A summary of the variables estimated in the field is presented in Table 1. Terrain slope in the plots averaged 10.7° and ranged from 0.7 to 25.7°. As can be observed, structural variables related to Aleppo pine forest show a remarkable variability. Inventoried trees range from smaller trees (heights of 3.7 m and stem diameters of 8.6 cm) to bigger ones (heights of 11.3 m and stem diameter of 28.3 cm). This implies a subsequent variability in biomass estimations from these variables, as can be seen in standard deviation values of aboveground biomass (AGB), Wr and TB.

Table 1. Summary of the field plot data (n=45).

	Min	Max	Range	Mean	Standard deviation
Slope (degrees)	0.7	25.7	25.0	10.7	7.2
Tree height (m)	3.7	11.3	7.6	7.1	2.0
dbh (cm)	8.6	28.3	19.7	15.2	4.7
AGB: Aboveground biomass (kg/ha)	989.2	112,076.4	111,087.2	37,462.6	27,879.5
Wr: Belowground biomass (kg/ha)	423.4	32,701.5	32,278.2	13,237.4	7,435.3
TB: Total biomass (kg/ha)	1,412.6	144,777.9	143,365.3	50,700.0	35,123.8
Carbon content (kg/ha)	704.9	72,244.2	71,539.3	25,299.3	17,526.8

Table 2 shows the Spearman correlation coefficients obtained for each of the ALS-derived metrics. TB is strongly correlated to the upper ALS height percentiles, particularly with the P₂₅ to P₆₀ (0.95 in all of them), being positive in all cases. The greater the height of the trees within the plot the greater the TB value as the stems and branches are more developed. The same interpretation presents the canopy height metrics, such maximum (0.88), mean (0.95) and mode elevation (0.92). In the case of metrics related to the variability of canopy height, the skewness of the height distribution of the returns, which describes the degree of concentration of heights around high or low values, i.e.,

provides some indication of how asymmetric the point distribution is, presents an inverse Spearman's correlation coefficient (-0.83). On the contrary, the standard deviation and variance, both with *Rho* values of 0.71, show the dispersion of point height values data around the mean height. Thus, the higher the TB the more dispersion of the data concerning the average height value because the trees are bigger. Finally, metrics related to canopy density correlate strongly with TB, particularly the percentage of first returns above mean height (0.88) and the ratios of all returns above mean and 1 m (0.88 and 0.84, respectively). The interpretation of such metrics is similar, as ratios express the proportion of returns belonging either to the canopy surface (first returns) or the whole canopy (all returns).

Table 2. Correlation coefficients (*Rho*) describing the strength of linear relationships between plot-derived TB and ALS-derived metrics. ** and * denote significance (P-value < 0.01 and 0.05, respectively); ^{ns} denote not significant (P-value > 0.05).

Variable	<i>Rho</i>	Variable	<i>Rho</i>
Canopy height percentiles metrics		Canopy height variability metrics	
P ₀₁	0.43**	Elev. standard deviation (SD)	0.71**
P ₀₅	0.67**	Elev. variance	0.71**
P ₁₀	0.85**	Elev. coefficient of variation (CV)	-0.37**
P ₂₀	0.94**	Elev. interquartile range (IQ)	0.54**
P ₂₅	0.95**	Elev. skewness	-0.83**
P ₃₀	0.95**	Elev. kurtosis	0.50**
P ₄₀	0.95**	Canopy density metrics	
P ₅₀	0.95**	% first returns above 1 m	0.83**
P ₆₀	0.95**	% all returns above 1 m	0.65**
P ₇₀	0.94**	% first returns above mean height	0.88**
P ₇₅	0.94**	% first returns above mode height	0.40**
P ₈₀	0.94**	% all returns above mean height	0.79**
P ₉₀	0.93**	% all returns above mode height	0.18 ^{ns}
P ₉₅	0.90**	(All returns above 1 m) / (Total first returns) · 100	0.84**
P ₉₉	0.89**		
Canopy height metrics			
Elev. minimum	0.10 ^{ns}	(All returns above mean) / (Total first returns) · 100	0.88**
Elev. maximum	0.88**		
Elev. mean	0.95**	(All returns above mode) / (Total first returns) · 100	0.40**
Elev. mode	0.92**		

Table 3 shows the summary of the selected model for estimating TB that includes three types of ALS metrics: i) the maximum elevation, ii) the kurtosis of the height distribution of the returns, and iii) the percentage of all returns above 1 m. As can be seen in Figure 1 and Table 2, the relationship between TB and these variables is direct. The higher the value of these metrics, the larger the amount of points returning from trees, implying an increase in timber volume per hectare and in plant material.

As shown in Table 3, the model presented a R^2 of 0.89, a RMSE of 11,327.69 and a RMSEcv of 11,418.11 kg/ha. The bias value of 0.12 evidences a slightly model overestimation.

Table 3. Model summary for estimating TB.

Dependent variable	Predictive model	Fitting phase			Cross-validation		
		R ²	RMSE	Bias	R ²	RMSEcv	Bias
TB (kg/ha)	-70277.86+8902.59·Elev. maximum +8232.54·Elev. kurtosis + 563.10·% all returns above 1 m	0.89	11,327.69	0.00	0.89	11,418.11	0.12

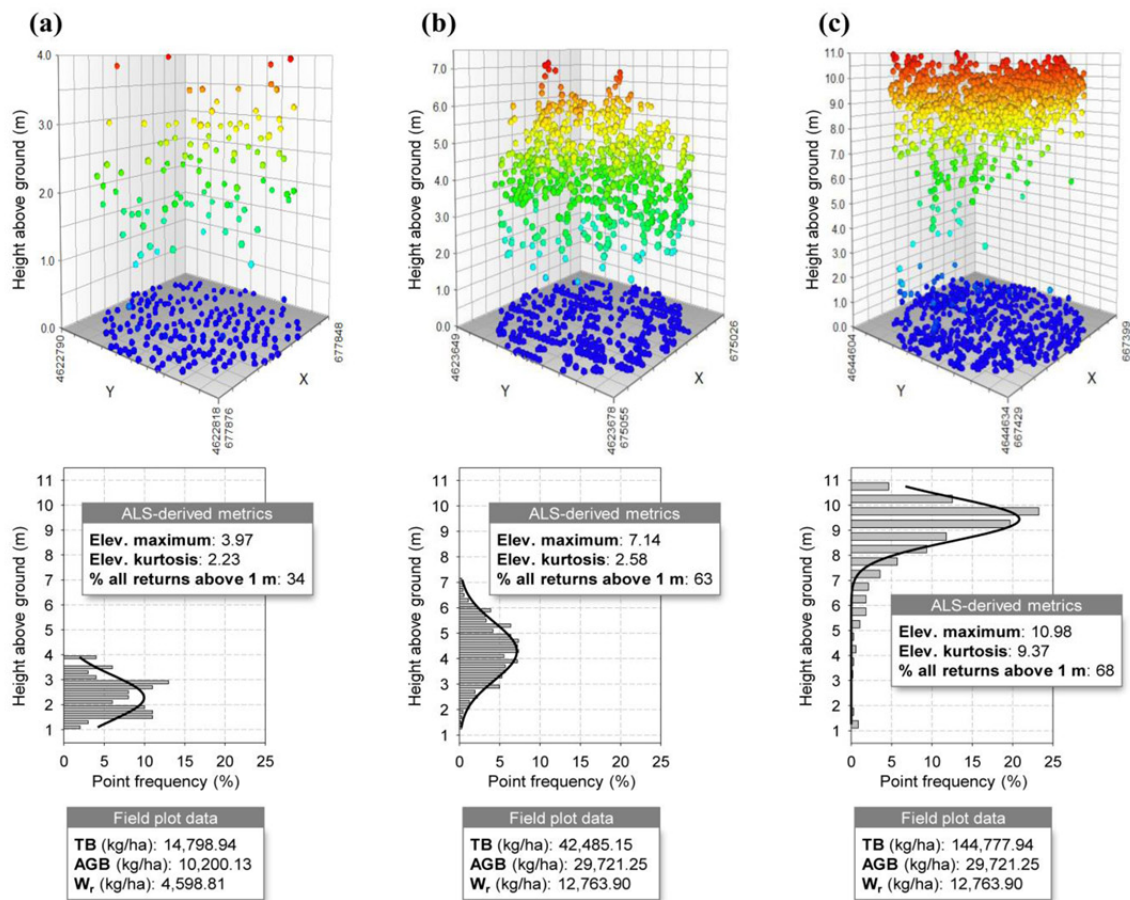
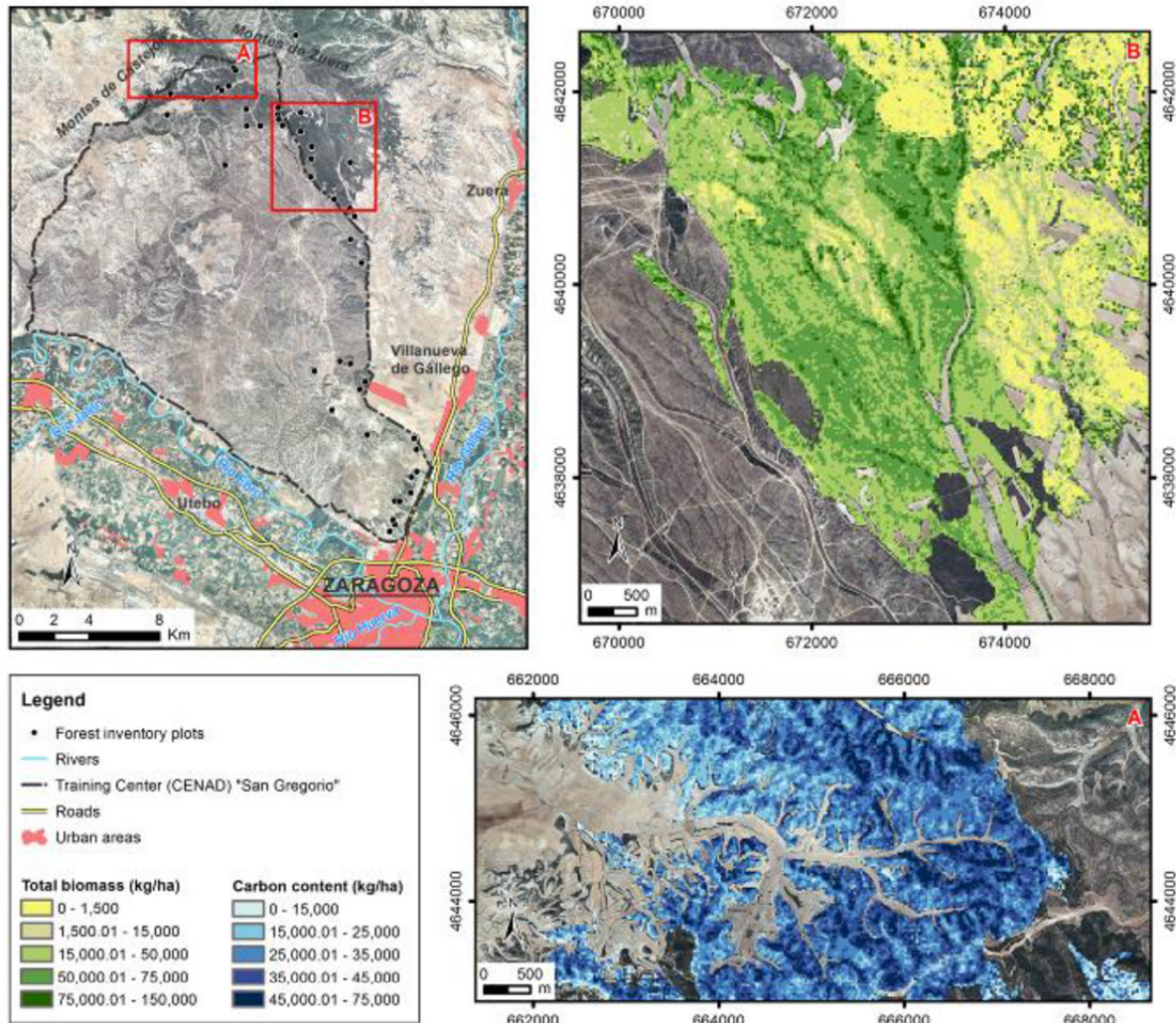
Figure 1. Metrics associated to the vertical distribution of ALS returns in three selected field plots representative of the study area: (a) Low height, (b) medium height and (c) great height Aleppo pines.

Figure 2 shows the TB and carbon content mapping after implementation of the prediction model in a GIS environment.

The results of this study indicate that LiDAR-PNOA data can be used to estimate the TB in monospecific *Pinus halepensis* Mill. forest stands. The good fit of the model obtained reveals, not only the suitability of the methodology, but also the correlation between the statistics obtained from ALS point cloud and the h and dbh measured in the field. Although the RMSE above 11,000 kg/ha may seem high, it should be noted that the inventory data covered a wide range of values as can be observed in Table 1.

Figure 2. Map of the TB and carbon content in two selected areas of the study site. (a) Heterogeneous Aleppo pine forest with lower stand density. (b) Includes two distinct homogeneous areas, differentiated by stand density and canopy height.



It is important to notice that both, the R^2 and RMSE value match those obtained by other authors, such as González-Ferreiro *et al.* [11]. In contrast to other studies, such as García *et al.* [25], it was not necessary to perform a logarithmic transformation of the original variables in order to achieve a better fit of the model or to meet the assumptions of the linear regression model.

ALS metrics included in the model behave logically as can be observed in Figure 1. In this regard, it is common that prediction models include ALS-derived metrics directly related to the canopy height, variables characterizing the variability, dispersion and shape of the height distribution of returns, as is the case of kurtosis, and variables related to the canopy density and tree cover, as the percentage of all returns above 1 m.

On the other hand, the no contemporaneity of the ALS data acquisition and the field campaign was considered irrelevant, as no significant changes in forest structure occurred in that time lapse.

Furthermore, the low point density of the data did not affect the existence of high and significant correlations between TB derived from field data and ALS-derived metrics.

In the future, it would be advisable to analyze the possibility of estimating forest biomass disaggregated in different fractions (stem, branches, leaves and roots) in order to model the carbon storage in detail and improve forest stand management. Besides, it would be desirable to study the usefulness of the PNOA-LiDAR data to characterize other forest species in large areas, as it could increase the accuracy of the results and reduce the costs with respect to traditional inventories.

4. Conclusions

Active remote sensing with ALS technology brings a new perspective to forest inventories providing 3-D information of the vegetation structure, as well as an improvement in the accuracy of the results and a reduction in costs. In this regard, the benefit of estimating the TB is even bigger, as with this methodology systematic sampling destructive procedures are not necessary. This work has proved the usefulness to estimate TB and carbon content in monospecific Aleppo pine stands representative of the continental Mediterranean climate. In the light of the results obtained, the adopted methodology was suitable. Our results match those achieved in other types of forests. Considering that PNOA project will presumably produce more LiDAR products in the next years, it would be desirable to evaluate the adequacy of the generated model to future captures of data and to develop new models of biomass components as well as in other types of Mediterranean forests.

Acknowledgments

This work has been financed by the Government of Aragón, Department of Science, Technology and University (FPI Grant BOA 30, 11/02/2011) and supported by the Research Project of Centro Universitario de la Defensa de Zaragoza (Project No: 2013-04). The ALS data were provided by the Spatial Information Centre of Aragón. The authors are grateful to the Training Center (CENAD) “San Gregorio” for its assistance in the field and for the invaluable technical support.

Author Contributions

M.T. Lamelas had the original idea for the study. Antonio L. Montealegre, M.T. Lamelas, A. García-Martín, J. de la Riva and F. Escribano conducted the fieldwork campaign. Antonio L. Montealegre, M.T. Lamelas, J. de la Riva and A. García-Martín developed the methodology and performed the analysis and Antonio L. Montealegre wrote the manuscript, incorporating suggestions from all co-authors, who approved the final manuscript.

Conflicts of Interest

The authors declare no conflict of interest.

References

1. Maltamo, M.; Næsset, E.; Vauhkonen, J. *Forestry Applications of Airborne Laser Scanning: Concepts and Case Studies*. Springer London, UK, 2014; p. 464.

2. Chuvieco, E. *Earth Observation of Wildland Fires in Mediterranean Ecosystems*. Springer Alcalá de Henares, Spain, 2009; p. 257.
3. Le Toan, T.; Beaudoin, A.; Riou, J.; Guyon, D. Relating forest biomass to SAR data. *Geoscience and Remote Sensing, IEEE Transactions on* **1992**, 30 (2), 403-411.
4. Leboeuf, A.; Fournier, R.A.; Luther, J.E.; Beaudoin, A.; Guindon, L. Forest attribute estimation of northeastern Canadian forests using QuickBird imagery and a shadow fraction method. *Forest Ecology and Management* **2012**, 266 (0), 66-74.
5. Tanase, M.; de la Riva, J.; Santoro, M.; Pérez-Cabello, F.; Kasischke, E. Sensitivity of SAR data to post-fire forest regrowth in Mediterranean and boreal forests. *Remote Sensing of Environment* **2011**, 115 (8), 2075-2085.
6. Vosselman, G.; Maas, H.-G. *Airborne and Terrestrial Laser Scanning*. Whittles Publishing Dunbeath, UK, 2010; p. 320.
7. Lefsky, M.A.; Cohen, W.B.; Parker, G.G.; Harding, D.J. Lidar Remote Sensing for Ecosystem Studies. *BioScience* **2002**, 52 (1), 19-30.
8. Kraus, K.; Pfeifer, N. Determination of terrain models in wooded areas with airborne laser scanner data. *ISPRS Journal of Photogrammetry and Remote Sensing* **1998**, 53 (4), 193-203.
9. Montealegre, A.L.; Lamelas, M.T.; Tanase, M.; de la Riva, J. Forest Fire Severity Assessment Using ALS Data in a Mediterranean Environment. *Remote Sensing* **2014**, 6 (5), 4240-4265.
10. García, D.; Godino, M.; Mauro, F. *Lidar: Aplicación Práctica Al Inventario Forestal*. Editorial Académica Española Lexington, USA, 2012; p. 196.
11. Gonzalez-Ferreiro, E.; Dieguez-Aranda, U.; Miranda, D. Estimation of stand variables in *Pinus radiata* D. Don plantations using different LiDAR pulse densities. *Forestry* **2012**, 85, 281 - 292.
12. Means, J.; Acker, S.; Fitt, B.; Renslow, M.; Emerson, L.; Hendrix, C. Predicting forest stand characteristics with airborne scanning LiDAR. *Photogrammetric Engineering and Remote Sensing* **2000**, 66 (11), 1367 - 1371.
13. Næsset, E. Predicting forest stand characteristics with airborne scanning laser using a practical two-stage procedure and field data. *Remote Sensing of Environment* **2002**, 80 (1), 88-99.
14. Bouvier, M.; Durrieu, S.; Fournier, R.A.; Renaud, J.-P. Generalizing predictive models of forest inventory attributes using an area-based approach with airborne LiDAR data. *Remote Sensing of Environment* **2015**, 156 (0), 322-334.
15. Magnussen, S.; Eggermont, P.; LaRiccia, V. N. Recovering Tree Heights from Airborne Laser Scanner Data. *Forest Science* **1999**, 45 (3), 407-422.
16. Næsset, E.; Bjerknes, K.-O. Estimating tree heights and number of stems in young forest stands using airborne laser scanner data. *Remote Sensing of Environment* **2001**, 78 (3), 328-340.
17. Holmgren, J.; Nilsson, M.; Olsson, H. Estimation of Tree Height and Stem Volume on Plots Using Airborne Laser Scanning. *Forest Science* **2003**, 49 (3), 419-428.
18. Hall, S.A.; Burke, I.C.; Box, D.O.; Kaufmann, M.R.; Stoker, J.M. Estimating stand structure using discrete-return lidar: an example from low density, fire prone ponderosa pine forests. *Forest Ecology and Management* **2005**, 208 (1-3), 189-209.
19. Coops, N.; Hilker, T.; Wulder, M.; St-Onge, B.; Newnham, G.; Siggins, A.; Trofymow, J. Estimating canopy structure of Douglas-fir forest stands from discrete-return LiDAR. *Trees-Structure and Function* **2007**, 21, 295 - 310.

20. Næsset, E.; Gobakken, T.; Holmgren, J.; Hyypä, H.; Hyypä, J.; Maltamo, M.; Nilsson, M.; Olsson, H.; Persson, Å.; Söderman, U. Laser scanning of forest resources: the nordic experience. *Scandinavian Journal of Forest Research* **2004**, *19* (6), 482-499.
21. Treitz, P.; Lim, K.; Woods, M.; Pitt, D.; Nesbitt, D.; Etheridge, D. LiDAR Sampling Density for Forest Resource Inventories in Ontario, Canada. *Remote Sensing* **2012**, *4* (4), 830-848.
22. Watt, M.; Meredith, A.; Watt, P.; Gunn, A. Use of LiDAR to estimate stand characteristics for thinning operations in young Douglas-fir plantations. *New Zealand Journal of Forestry Science* **2013**, *43* (1), 18.
23. Zhao, K.; Popescu, S.; Nelson, R. Lidar remote sensing of forest biomass: A scale-invariant estimation approach using airborne lasers. *Remote Sensing of Environment* **2009**, *113* (1), 182-196.
24. Næsset, E.; Gobakken, T. Estimation of above- and below-ground biomass across regions of the boreal forest zone using airborne laser. *Remote Sensing of Environment* **2008**, *112* (6), 3079-3090.
25. García, M.; Riaño, D.; Chuvieco, E.; Danson, F.M. Estimating biomass carbon stocks for a Mediterranean forest in central Spain using LiDAR height and intensity data. *Remote Sensing of Environment* **2010**, *114* (4), 816-830.
26. Vosselmann, G.; Maas, H.G. *Airborne and terrestrial laser scanning*. Whittles Publishing Dunbeath, UK, 2010; p. 318.
27. Vicente-Serrano, S.M.; Lasanta, T.; Gracia, C. Aridification determines changes in forest growth in *Pinus halepensis* forests under semiarid Mediterranean climate conditions. *Agricultural and Forest Meteorology* **2010**, *150* (4), 614-628.
28. Evans, J.S.; Hudak, A.T. A Multiscale Curvature Algorithm for Classifying Discrete Return LiDAR in Forested Environments. *Geoscience and Remote Sensing, IEEE Transactions on* **2007**, *45* (4), 1029-1038.
29. Renslow, M. *Manual of Airborne Topographic Lidar*. The American Society for Photogrammetry and Remote Sensing Bethesda, USA, 2013; p. 504.
30. McGaughey, R. *FUSION/LDV: Software for LIDAR Data Analysis and Visualization*, US Department of Agriculture, Forest Service, Pacific Northwest Research Station Seattle, USA, 2009; p. 175.
31. Næsset, E.; Okland, T. Estimating tree height and tree crown properties using airborne scanning laser in a boreal nature reserve. *Remote Sensing of Environment* **2002**, *79* (1), 105 - 115.
32. Ruiz-Peinado, R.; del Rio, M.; Montero, G. New models for estimating the carbon sink capacity of Spanish softwood species. *Forest Systems* **2011**, *20* (1), 176-188.
33. Hair, J.F.; Anderson, R.E.; Tatham, R.L.; Black, W.C. *Análisis multivariante*. 5th ed.; Prentice Hall Iberia Madrid, Spain, 1999; p. 799.
34. Chen, Q.; Gong, P.; Baldocchi, D.; Tian, Y.Q. Estimating Basal Area and Stem Volume for Individual Trees from Lidar Data. *Photogrammetric Engineering & Remote Sensing* **2007**, *73* (12), 1355-1365.
35. Andersen, H.-E.; McGaughey, R.J.; Reutebuch, S.E., Estimating forest canopy fuel parameters using LIDAR data. *Remote Sensing of Environment* **2005**, *94* (4), 441-449.

36. Montero, G.; Ruiz-Peinado, R.; Muñoz, M. *Producción de biomasa y fijación de CO₂ por los bosques españoles*. Instituto Nacional de Investigación y Tecnología Agraria y Alimentaria, Ministerio de Educación y Ciencia Torrejón de Ardoz, Spain, 2005; p. 270.

© 2015 by the authors; licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution license (<http://creativecommons.org/licenses/by/4.0/>).

3.5 Modelo logístico para la evaluación de la severidad post-incendio

RESUMEN

El fuego ha sido un elemento modelador del paisaje y de los ecosistemas mediterráneos a lo largo de la historia. Además, este es un fenómeno natural que está cada vez más influenciado por la acción antrópica. Los pinares mediterráneos españoles se ven afectados por incendios forestales con diferente frecuencia, intensidad y severidad, dando lugar a consecuencias socio-ecológicas de diversa índole. Entre ellas, cabe destacar el grado de impacto sufrido por la vegetación tras el fuego. Para su valoración, hay que recurrir a estimaciones lo más precisas posibles de la severidad, la cual podrá ayudar a los gestores del bosque a decidir qué estrategia es más adecuada para mitigar el efecto del fuego. Tradicionalmente, se ha hecho uso de índices espectrales derivados de imágenes de teledetección óptica multiespectral en combinación con parcelas de muestreo en el terreno, para relacionar los cambios radiométricos detectados con los efectos reales. Sin embargo, la información tridimensional proporcionada por los sistemas activos LiDAR aeroportados ha sido poco explorada con este propósito.

El objetivo de esta investigación es la estimación de la severidad post-incendio, relacionando un conjunto de variables independientes derivadas de las nubes de puntos del vuelo LiDAR del Plan Nacional de Ortofotografía Aérea (PNOA), con datos de campo basados en el índice CBI (*Composite Burn Index*) y recogidos en cuatro incendios localizados en Aragón. Se ha ajustado un modelo de regresión logística, que tras la validación, ha reportado una precisión del 85,5%, siendo las variables "*canopy relief ratio*" y el porcentaje de puntos por encima de 1 m de altura sobre el terreno, las incluidas en dicho modelo. Por otro lado, se ha realizado una comparativa de los resultados con índices espectrales derivados de imágenes Landsat Thematic Mapper.

Article

Forest Fire Severity Assessment Using ALS Data in a Mediterranean Environment

Antonio Luis Montealegre ^{1,*}, María Teresa Lamelas ^{1,2}, Mihai A. Tanase ³ and Juan de la Riva ¹

¹ GEOFOREST Group, IUCA, Department of Geography, University of Zaragoza, Zaragoza 50009, Spain; E-Mails: tlamelas@unizar.es (M.T.L.); delariva@unizar.es (J.R.)

² Centro Universitario de la Defensa de Zaragoza, Zaragoza 50090, Spain

³ Department of Infrastructure Engineering, University of Melbourne, Melbourne, VIC 3052, Australia; E-Mail: mihai@tma.ro

* Author to whom correspondence should be addressed; E-Mail: monteale@unizar.es; Tel.: +34-876-554-058; Fax: +34-976-761-506.

Received: 2 January 2014; in revised form: 8 April 2014 / Accepted: 15 April 2014 /

Published: 8 May 2014

Abstract: Mediterranean pine forests in Spain experience wildland fire events with different frequencies, intensities, and severities which result in diverse socio-ecological consequences. In order to predict fire severity, spectral indices derived from remotely sensed images have been used extensively. Such spectral indices are usually used in combination with ground sampling to relate detected radiometric changes to actual fire effects. However, the potential of the tridimensional information captured by Airborne Laser Scanners (ALS) to severity mapping has been less explored. With the objective of addressing this question, in this paper, explanatory variables extracted from ALS point clouds are related to field estimations of the Composite Burn Index collected in four fires located in Aragón (Spain). Logistic regression models were developed and statistically tested and validated to map fire severity with up to 85.5% accuracy. The canopy relief ratio and the percentage of all returns above one meter height were the most significant variables and were therefore used to create a continuous map of severity levels.

Keywords: fire severity; composite burn index; Airborne Laser Scanners (ALS); Mediterranean pine forest; logistic regression

1. Introduction

Fire is a global phenomenon [1] with 200–500 million hectares being burned annually [2]. Fires affect large areas in a variety of biomes from tropical to boreal being more widespread than any other natural disturbance [2,3]. In some ecosystems, e.g., savannas and grasslands, fire plays an ecologically significant role in biogeochemical cycles and disturbance dynamics. In other ecosystems, fire may lead to the destruction of forests or to long-term site degradation.

In Mediterranean Europe, fire is a major hazard with an average of 45,000 fires being recorded yearly [4]. Therefore, fire constitutes one of the main factors determining the current forest landscape [5], with summertime wildfires causing extensive ecological and economic losses [6] over hundreds of thousands of hectares of forests, shrub lands and grasslands. In Spain, although the last decade, 2001–2010, has seen a reduction in the number of fire incidents (averaging 17.127 per year), the occurrence of large fires (>100 ha) increased. In 2012, 64% of the total area affected by fires was burned in a large fire [7]. Even though Mediterranean species are adapted to natural fire regimes, changes in land use and the cumulative effects of anthropogenic disturbances make vegetation communities more vulnerable to high-intensity wildfires. In addition, climate change forecasts imply an increasing vulnerability of our forests to large-scale fires in the future [8]. In response, scientists and fire managers require the most accurate information available regarding the impact of fire on the environment. The way fire is distributed throughout the affected area is essential for short-term mitigation and restoration treatments, as well as for vegetation recovery monitoring, wildlife studies, soil and hydrologic changes, as well as various ecological processes [9].

In most cases, fire severity and burn severity are used interchangeably to describe fire effects on ecosystems in terms of biophysical alteration (e.g., blackening or scorching of tree-leaves, amount of fuel consumed, depth of burn, soil exposure, *etc.*). However, these two terms may imply different meanings. Fire severity can be defined as a measure of the immediate fire impact on the environment such as tree mortality and the loss of biomass in the forms of vegetation and soil organic material, while burn severity can be defined as the degree of ecological change caused by fire [9,10]. The Composite Burn Index (CBI) was developed by Key and Benson [11] within the framework of the Fire Effects Monitoring and Inventory Protocol (FIREMON) project for pine forests in western USA, to sample these changes and to summarize the general effects of a fire at a given plot. The CBI has been used and adjusted [9] to a wide range of environments, from Mediterranean to boreal [12–19].

Extensive field surveys are often costly and labor intensive, making the use of remotely sensed data indispensable for fire severity analysis. The removal of vegetation, the exposure of soil, and changes in soil and vegetation moisture content as a result of differences in burn severity, imply changes in the near and short-wave infrared regions of the electromagnetic spectrum that can be detected by multispectral remote sensing devices. Remote sensing products are not only more cost-effective when compared to ground-based methods, but also provide information even in inaccessible areas shortly after fire events [9].

The estimation of burn severity from satellite data has been accomplished using per band reflectance or vegetation spectral indices derived from optical sensors of different spatial resolutions, from 1 km (e.g., Advanced Very High Resolution Radiometer—AVHRR) down to 2 m (e.g., QuickBird), with Landsat satellite images (30 m pixel) being the most commonly used [20]. However, estimations from

passive satellite sensors are less sensitive to changes in forest structure [21] than active sensors, such as Radio Detection and Ranging (radar) [22,23] and Airborne Laser Scanner (ALS) [24].

ALS has already been adopted and accepted as a valuable tool in forestry applications due to the three-dimensional nature of the data. The laser pulse penetration characteristics and the multi-return recording capabilities enable an accurate geometry-related approach in forestry, and by extension in wildland fire management [25]. Research focused on this technology has moved from validation of ALS measurement effectiveness for forest studies (e.g., [26,27]), to estimation of continuous variables such as basal area and biomass (e.g., [28,29]), as well as to forest structure analysis over large areas (e.g., [30–33]) and the quantification and assessment of canopy gaps and patches within forests (e.g., [24,30,34]). Increasingly, ALS data is being used with the aim of estimating fuel parameters, such as crown bulk density or height to live crown, which are commonly introduced as inputs in fire behavior models (e.g., [28,35–40]).

In this context, the present study aims to develop a methodology that combines field data, represented by CBI estimations, and information derived from multiple-return ALS data, to assess fire severity for forest management purposes using multivariate analysis, specifically logistic regression. The following questions were addressed:

1. What is the empirical relationship between field-measurements of burn severity and ALS data variables derived from point height distribution and number of returns?
2. Which is the best model that will allow for the estimation of fire severity continuously?

2. Material and Methods

In this study, fire severity was evaluated over the area covered by four wildfire events in 2008–2009. In order to characterize the spatial variability of the impact caused by fire, a wide range of variables were generated using ALS data. Then, a multivariate statistical analysis was carried out in order to evaluate which ALS-derived measurements present the strongest relationship with ground-assessed severity. The last but not least purpose was to map the spatial distribution of fire severity. Validation procedures were adopted to evaluate the quality, reliability, robustness, and degree of fitting of the results.

2.1. Study Area

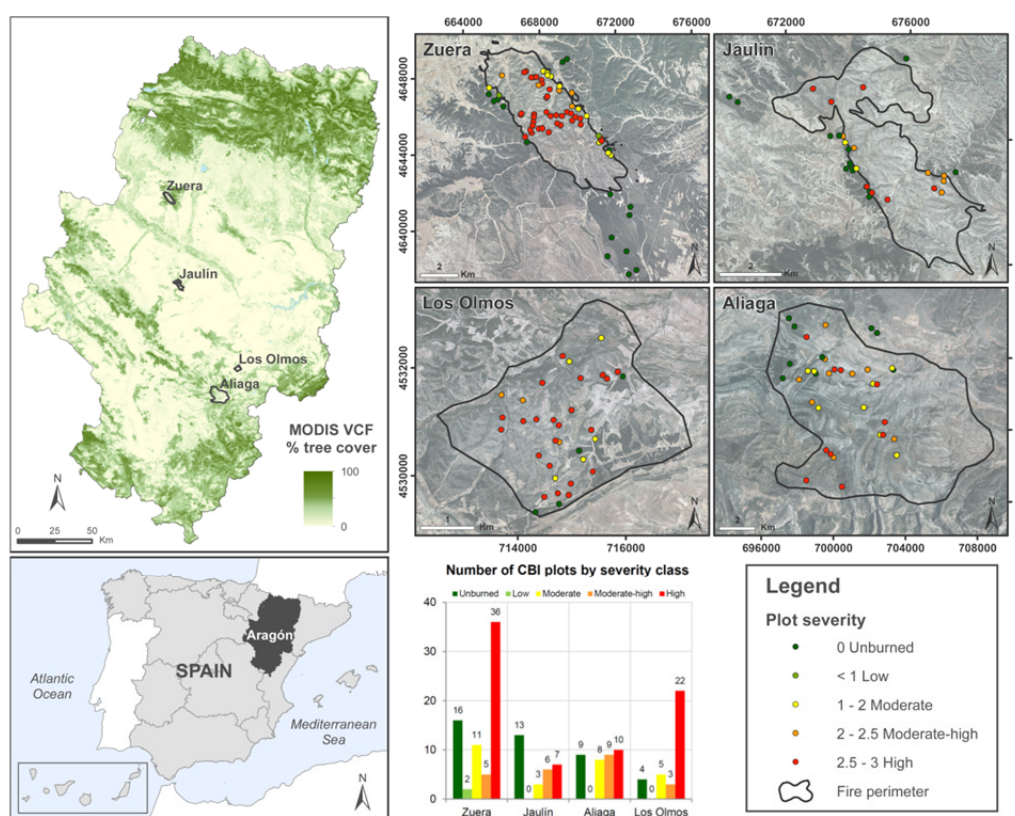
The Autonomous Region of Aragón, located in northeastern Spain (Figure 1), stretches from the Pyrenees range, in the north, to the Iberian range, in the south. Aragón is the northernmost semi-arid region in Europe, although it is crossed by the Ebro River Valley at its center. The geological diversity of this area results in high topographic variation: The Pyrenees rise to more than 3000 m.a.s.l. and extend their foothills, the pre-Pyrenees, southward to the Ebro basin, gradually decreasing in elevation up to 100 m.a.s.l.

In the central sector of the Ebro Valley, annual precipitation is low, averaging 350 mm and mostly occurring in autumn and spring. Its Mediterranean continental climate [41] presents cold winters, with monthly mean temperature about 7 °C, and hot, dry summers, with mean temperatures about 24 °C. However, altitude and topographic barriers modify this general pattern of extreme aridity at the edge of

the Ebro Basin by increasing humidity and diminishing temperatures, while keeping the same pattern of seasonal rain and drought as the center of the Ebro valley.

Approximately, 50% of the 1.5 million hectares of forested areas in Aragón are composed of coniferous forest, 35% correspond to mixed forest, and the rest is covered by deciduous forest. Open forests, with less than 20% of tree cover, and shrublands occupy another one million hectares. Over the last few decades, the area covered by forest increased as a result of afforestation projects and agricultural abandonment. The main tree species are *Pinus halepensis* P. Mill., *Pinus sylvestris* L., *Pinus pinaster* Soland., non Ait., *Quercus ilex* L., and *Quercus pyrenaica* Willd. The species *Quercus coccifera* L., *Juniperus oxycedrus* L. subsp. *macrocarpa* (Sibth. & Sm.) Ball, and *Thymus vulgaris* L. are commonly encountered in shrublands [20].

Figure 1. Study area and Composite Burn Index (CBI) field plots location. MODIS vegetation fractional cover (VFC) images and high spatial resolution ortophotography (PNOA-2012) are used as backdrop.



Wildfires have significantly increased in the last twenty years, both in number and total area. The increase was related not only to weather conditions, but also to human activity. In Aragón, approximately 5000 ha are affected annually by fire. Almost 3000 ha of this burned land are forested areas. In exceptional years such as 1994 and 2009, the affected area reached up to 30000 ha. In 2008 and 2009, eight large fires (>100 ha) were declared in Aragón and the total affected area was around 23,000 ha, most of which was burned in 2009. All of the fires affected pine forests after a period of extremely dry

and hot weather. In six of those fires the cause of ignition was lightning while the remaining two were ignited due to anthropogenic activities [20]. Four of these fires (Figure 1) were analyzed in this study as a consequence of their significant environmental impact due to their large extension (Table 1).

Table 1. Fire events and the corresponding field plots analyzed.

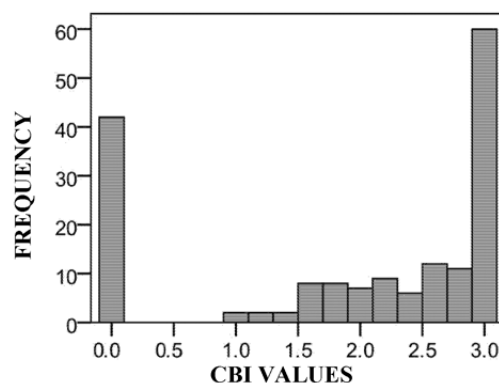
Fire	Date	Cause	Burned Area (ha)	CBI Plots
Zuera	August 2008	Human	2200	70
Jaulín	July 2009	Lightning	1800	29
Aliaga	July 2009	Lightning	9000	36
Los Olmos	July 2009	Lightning	500	34
Total			13,500	169

2.2. Field Data Collection

Field data were collected within two months after each fire utilizing the Composite Burn Index (CBI) field protocol, which takes into account the visible and averaged burn severity condition found in a plot [11]. Circular plots of 30 m diameter were laid out in areas of homogeneous fire effects using a Trimble GeoExplorer GPS with a positioning error at the center of the plot lower than 1 m after differential correction. Digital photos were taken from the center of each plot to the four cardinal directions to provide qualitative information about vegetation structure and soil conditions. The CBI assessment is a somewhat subjective estimation of the entire averaged burn severity across five forest strata: substrates, herbaceous vegetation, large shrubs and small trees, intermediate trees, and dominant and co-dominant canopy trees. The CBI for the whole plot area is based on synoptic scores ranging from 0 (not-burned) to 3 (completely burned) as described by Key and Benson [11]. To reduce any assessor bias, CBI assessments were conducted by the same two individuals.

Concerning fire effects in Aragón, low severity areas were characterized by scorched trees blackened at the base, charred to partially consumed litter while the understory layer was affected in a patchy pattern. The foliage and smaller twigs of the understory were partially to completely consumed, while branches were mostly intact. The tree crowns remained largely unaffected. In moderate severity areas, part of the understory layer was preserved. The needles or leaves and the small stems of the short shrubs (<1 m height) were consumed, whereas the tall shrubs (>1 m height) presented scorched foliage. The tree crowns were partially scorched, which resulted in incomplete foliage loss. For moderate–high severity areas, the consumption of the understory layer increased, whereas the tree crowns were usually scorched, with the foliage being retained only partially. The loss of branches and twigs was small for the overstory. The understory layer lost the twigs and small branches, especially the small shrubs (*i.e.*, less than 1 m tall). In high severity areas, the combustion of the litter, understory layer, and tree crown elements (*i.e.*, needles and twigs) was complete. At the highest severities, small- and medium-sized branches were also consumed. For more detail on the field sampling, see Tanase *et al.* [20].

A selection of 169 from the initial 247 plots collected by Tanase *et al.* [20] was used in order to avoid areas where post-fire forest management activities, impacting canopy structure, had been carried out. For this reason, the 169 sampled plots do not represent proportionally the entire severity range (Figure 2). Table 2 shows the number of plots located in each severity range and how most of the plots were located in unburned or high severity areas.

Figure 2. Frequency histogram of the CBI values.**Table 2.** CBI field plots used as ground data by fire severity range and site ($N = 169$).

Fire	Severity Synoptic Scores				
	Unburned (CBI = 0)	Low (CBI ≤ 1)	Moderate (1 < CBI ≤ 2)	Moderate-High (2 < CBI ≤ 2.5)	High (CBI > 2.5)
Zuera	16	2	11	5	36
Jaulín	13	0	3	6	7
Aliaga	9	0	8	9	10
Los Olmos	4	0	5	3	22
Total	42	2	27	23	75

2.3. ALS Acquisition

The ALS data were provided by the Spanish National Plan for Aerial Orthophotography (PNOA) [42] and captured, in several surveys conducted between 1 August 2010 and 5 February 2011, using two distinctive airborne sensors, Leica ALS60 and ALS50-II laser scanners. Data were delivered in LAS binary file format containing X and Y coordinates (UTM Zone 30 ETRS 1989), ellipsoidal elevation Z (ETRS 1989), with up to four returns measured per pulse, and intensity values from a 1064 nm wavelength laser. The resulting ALS nominal point density was 0.5 points/m² with a vertical accuracy of less than 0.20 m. It should be noted that no difference exist between the ALS datasets in terms of sensing characteristics (*i.e.*, operating principles) because PNOA mission ensures homogeneous flight campaigns and products. Moreover, the seasonality did not affect the coherence of the acquisitions because the study site is covered entirely by evergreen vegetation.

ALS Processing

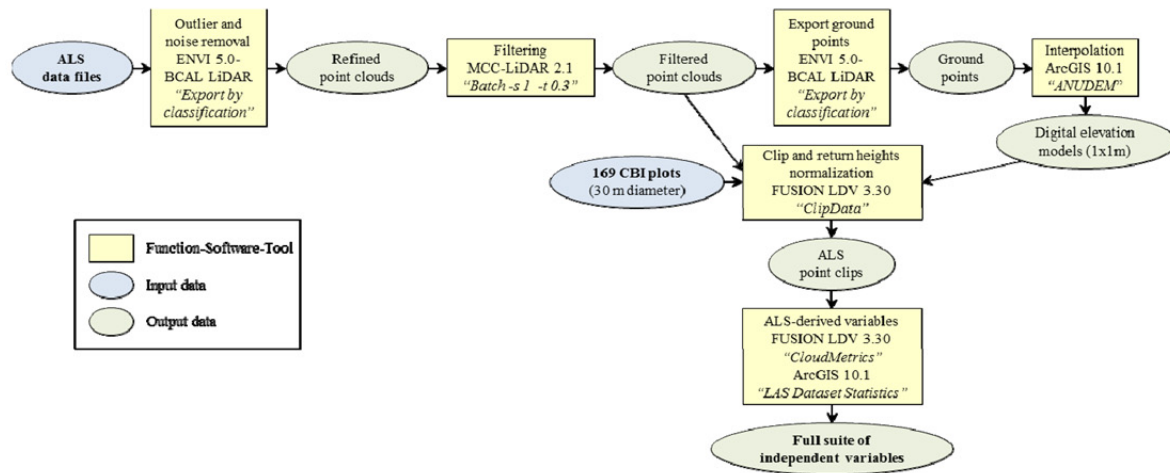
The ALS data set was provided in fifty-six 2 × 2 km tiles of raw data points which needed pre-processing before being filtered. Outliers and noise were removed with the open-source BCAL LiDAR module, developed by Idaho State University, Boise Center Aerospace Laboratory (BCAL), implemented in ENVI 5.0 (Boulder, CO, USA). Afterwards, the point cloud was filtered using MCC-LIDAR 2.1 command-line tool for processing discrete-return ALS data in forested environments [43]. The process classifies data points iteratively as ground or non-ground using the Multiscale Curvature Classification

(MCC) algorithm developed by Evans and Hudak [44]. According to Montealegre *et al.* [45], this classification algorithm, based on identifying non-ground points that exceed positive curvature thresholds across multiple scales, balances commission and omission errors in forestry applications. Subsequently, a high-resolution (1 m) bare earth digital elevation model (DEM), was created from the ground points using ANUDEM interpolation implemented in ArcGIS 10.1 software (ESRI, Redlands, CA, USA). ANUDEM uses a thin plate spline method where a smooth surface is constructed from the irregularly spaced ALS points [46]. Finally, the ALS point elevations were normalized by subtracting off the elevation of the underlying terrain, represented by the DEMs previously created, to obtain estimates of aboveground height for each return. This subtraction was performed using FUSION LDV 3.30 open source software [47], developed by the Remote Sensing Applications Center (USDA).

With the aim of getting statistical information about the laser returns in order to estimate fire severity, a full suite of independent variables (Table 3) commonly used in vegetation modelling [48] was generated with “ClipData” and “CloudMetrics” commands included in FUSION LDV 3.30 [49] and with “LAS Dataset” tool in ArcGIS 10.1. The computation was performed at plot level (30 m diameter). A flowchart of the ALS data processing is presented in Figure 3.

Table 3. ALS-derived variables description.

ALS variables	Description	ALS Variables	Description
%_num_of_ret_1; _2; _3; _4	% of pulses with one, two, three and four returns	Elev_maximum	Maximum height (m)
%_Class_Unassigned	% of object points	Elev_mean	Mean height (m)
%_Class_Ground	% of ground points	Elev_minimum	Minimum height (m)
%_Return_1; _2; _3	% of first, second and third returns	Elev_mode	Mode height (m)
Ratio_All_returns_1m	(All returns above 1 m)/ (Total first returns) × 100	Elev_P01;_P05;_P10 ... _P99	Percentiles (m) of point heights distribution
Ratio_All_returns_2m	(All returns above 2 m)/ (Total first returns) × 100	Elev_skewness	Skewness of point heights distribution
Ratio_All_returns_3m	(All returns above 3 m)/ (Total first returns) × 100	Elev_stddev	Standard deviation of point heights
Ratio_All_returns_mean	(All returns above mean height)/ (Total first returns) × 100	Elev_variance	Variance of point heights
Ratio_All_returns_mode	(All returns above mode height)/ (Total first returns) × 100	%_All_returns_1m; _2m; _3m	% all returns above 1, 2 and 3 m
Canopy relief ratio	((Mean height—Min height)/(Max height—Min height))	%_All_returns_mean	% all returns above mean value of point heights
Diff_Elev	Range of points elevation (m)	%_All_returns_mode	% all returns above mode value of point heights
Elev_AAD	Average Absolute Deviation of point heights	%_First_returns_1m; _2m; _3m	% first returns above 1, 2 and 3 m
Elev_IQ	Interquartile distance of point heights	%_First_returns_mean	% first returns above mean value of point heights
Elev_kurtosis	Kurtosis of point heights distribution	%_First_returns_mode	% first returns above mode value of point heights

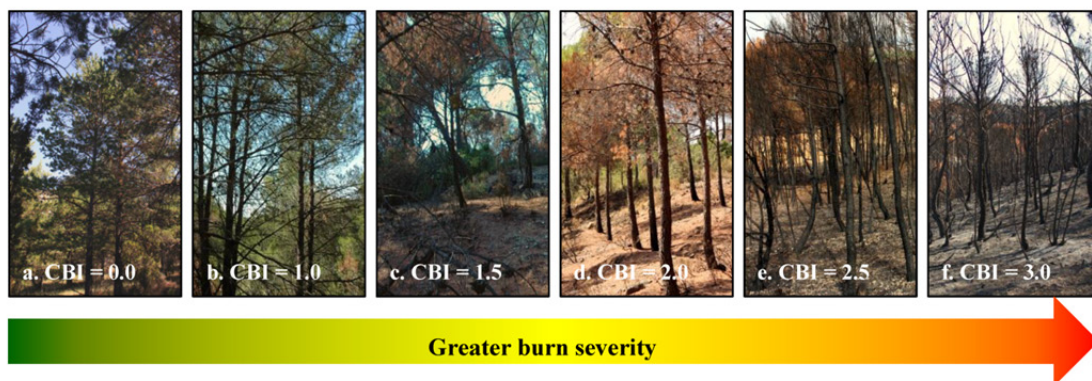
Figure 3. Flow diagram of the ALS data processing.

2.4. Data Analysis

The steps followed in the methodology used to evaluate the relationship between the dependent (CBI) and independent (ALS-derived) variables are presented in Table 4. The Spearman's rank correlation coefficient was used with the purpose of assessing the statistical dependence between plot severity and each ALS variable. A linear regression analysis was rejected as the assumption of normality of the CBI dataset was not met, even after transforming the variables. Thus, a logistic regression analysis, which requires splitting the CBI values into null or low severity and high severity, was selected. Therefore, the Kruskal–Wallis test was used for testing whether the mean values and distribution of the independent variables in both groups (null or low severity and high severity) are statistically different. This test is commonly used to select the most significant variables to be introduced in logistic regression analysis when it is not possible to assume normality and homoscedasticity of the samples [50]. As mentioned above, the dataset was split into two categories, null-low burn severity (CBI value ≤ 1.5) and high burn severity (CBI value > 1.5) to carry out this non-parametric method. The selection of the threshold value 1.5 was based on a qualitative analysis of the ALS point clouds and corresponding values of severity. The CBI method is based on a visual analysis of the magnitude of the environmental change with respect to the presumed pre-fire state. Therefore, given the information collected in the CBI forms such as the percentage of canopy mortality in big and intermediate trees, the depth of burn, percentage of the foliage altered, and the degree of change in vegetation strata, it was analyzed visually how these changes are reflected in the point clouds at plot level. Wildfire disturbance influences the horizontal and vertical structure of vegetation [34,51], which is recorded by ALS returns (Figure 4). However, as we only had post-fire information and it was not possible to estimate the structural change, it was assumed that the pre-fire structure of vegetation was homogeneous and similar to areas not affected by fire inside or near the fire perimeter (Figure 5a). Figure 5 depicts some examples of normalized point clouds in plots with different CBI values. The loss of canopy and the forest opening where logs remain without foliage are the most relevant indicators captured by the ALS. As can be observed in Figure 5, point clouds are relatively dense and homogeneous for CBI values up to 1.5 and from that value onwards, the amount of returns decrease

more significantly because the tree crowns were partially scorched with incomplete foliage loss. Therefore, we assumed that ALS was able to distinguish plots with CBI values under 1.5 (null and low severity) and above 1.5 (high severity). CBI values above 1.5 present a general trend of loss of vegetation structure in the field, which manifests itself in the ALS point clouds.

Figure 4. Partial views of different Composite Burn Index (CBI) plots depending on its burn severity value. Each photograph corresponds to the point cloud showed in the Figure 5.



As mentioned above, logistic regression analysis, one of the most frequently used multivariate statistical analysis, was selected due to the non-normal distribution of the variables [52,53]. It enables the creation of reliable predictive models from explanatory variables (either continuous or discrete). This technique, already used successfully to evaluate several phenomena—e.g., landslide hazard, water erosion, sinkhole susceptibility, human-caused ignition, *etc.*—evaluates the probability of an event occurrence (P) by estimating the possibility that a case will be classified into one of two mutually exclusive categories as opposed to the other category of the dependent dichotomous variable [54]. In this research, we used logistic regression to establish a functional relationship between the presence or not of fire impact within a plot in a binary-coded manner (1 = high burn severity; 0 = absence or low burn severity) and the set of the independent variables derived from ALS data. This can be written as:

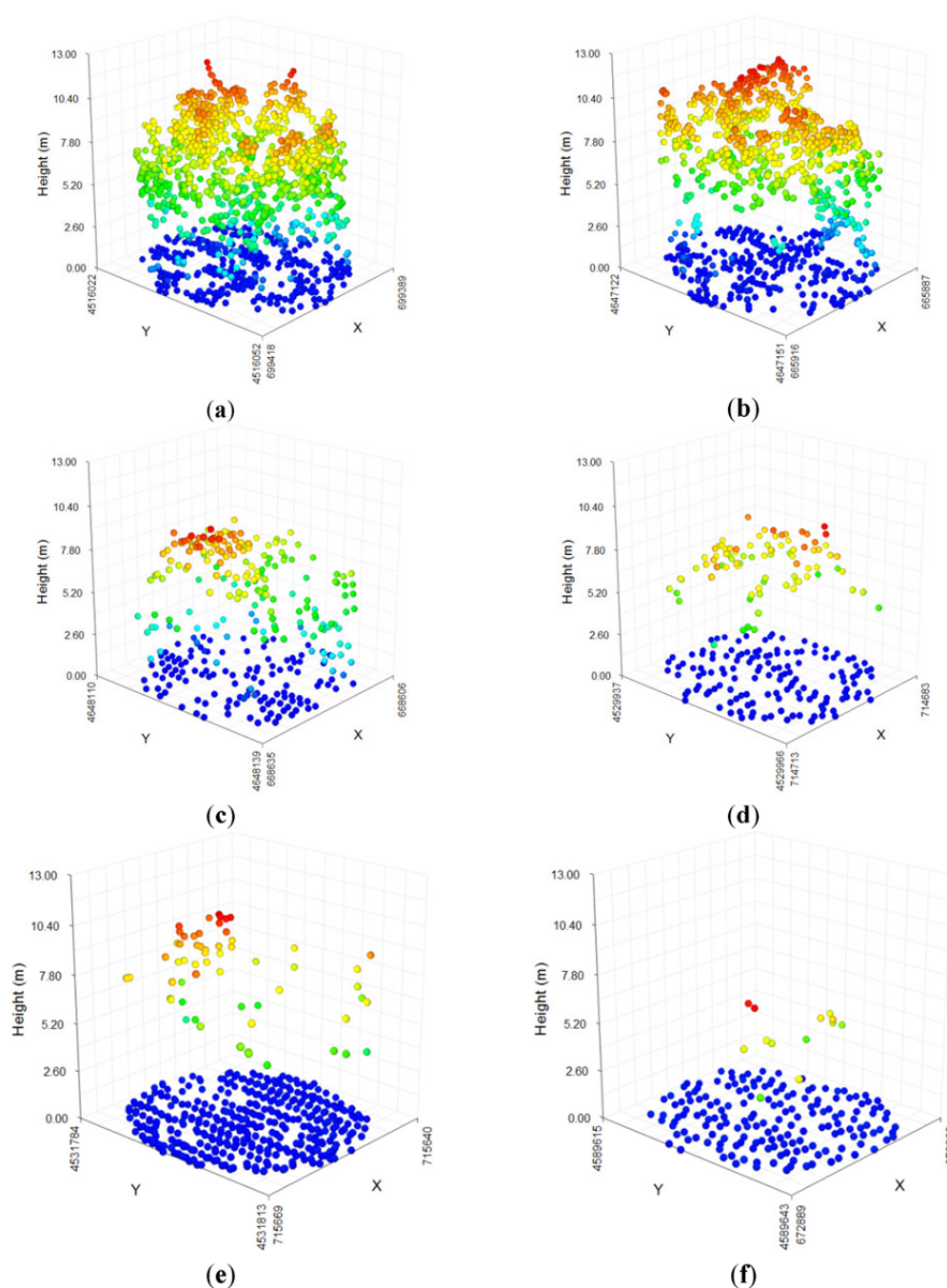
$$[P = \exp (B_0 + B_1X_1 + \dots + B_nX_n) / 1 + \exp (B_0 + B_1X_1 + \dots + B_nX_n)] \quad (1)$$

where P is the probability of occurrence, B_0 is the intercept of the model, X_n is a set of independent variables and B_n is a set of $n + 1$ parameters.

As the acceptance of a model requires the evaluation of its robustness to small changes of the input data, CBI data were split into two different samples. A random selection of 118 cases (70% of the sample) was used as a training subset, while the 51 unselected cases were used to test the accuracy of the model. Well-balanced samples were considered in terms of proportion of plots with high and low fire severity values. A forward stepwise (likelihood ratio) method was applied [55] to the training dataset in order to select the explanatory variables and to obtain several models. These models were evaluated by computing the Cox and Snell and the Nagelkerke pseudo- R^2 tests, the log likelihood ($-2LL$) statistic and the chi-square test in order to calculate how strong the relationship between the explanatory variables and the outcome was. The Hosmer and Lemeshow test is similar to a chi-square test, and it indicates the extent to which the model provides better fit than a null model with no

predictors, *i.e.*, how well the model fits the data, as in log-linear modelling. The significance was assessed individually for each independent variable incorporated in the model by means of the Wald test [50,54].

Figure 5. Examples of ALS point clouds at plot level and their correspondence with the CBI values. (a) CBI = 0.0; (b) CBI = 1.0; (c) CBI = 1.5; (d) CBI = 2.0; (e) CBI = 2.5; (f) CBI = 3.0.



The model was selected based on the principle of parsimony that implies a preference of models with a smaller number of parameters as each parameter introduced into the model adds some uncertainty to it, considering the significance level of individual predictor variables, as well as the usefulness of the model.

Logistic models are frequently used in a classification approach by selecting a given value of the response variable (the probability of high fire severity, in our case) and classifying all resulting values in one of two groups (1 or 0) according to it. The threshold value is normally the 0.5 probability, as usually the two sample groups are similar in size. However, for the case where the two groups are very dissimilar in size, the proportion of ones in the sample should be used instead of the 0.5 value [56,57]. Thus, value 1 was assigned to the predicted response when probability was greater than the cutoff value 0.6, specified by the proportion of CBI plots classified as 1 (CBI value > 1.5) in the observed values of the dependent variable.

The accuracy of selected logistic regression model was tested by cross tabulation associated with Cohen's Kappa index and the receiver operating characteristic (ROC) curves. The first is a statistical coefficient of inter-rater agreement that allows knowing if agreement is attributable to chance alone. Kappa index ranges generally from 0 to 1 (although negative numbers are possible). According to Landis and Koch [58], values of Kappa below 0.40 present poor agreement, from 0.40 to 0.75 are considered of good agreement, and above 0.75 generally reflect excellent agreement. In the case of ROC curves, the computed values of the area under the ROC curve (AUC) [59], provide a measure of goodness-of-fit of the logistic regression model based on the simultaneous measure of sensitivity (true positive) and specificity (true negative) for all possible cutoff points. Sensitivity is calculated as the fraction of CBI plots that were correctly classified as high burn severity, while specificity is derived from the proportion of CBI plots that were properly categorized as low burn severity. The closeness of the ROC curve to the upper left corner ($AUC = 1$) indicates a high accuracy of the model (*i.e.*, a correct discrimination between positive and negative cases). In relation to the computed AUC value, Hosmer and Lemeshow [52] classify a predictive performance as acceptable ($AUC > 0.7$), excellent ($AUC > 0.8$) or outstanding ($AUC > 0.9$).

Finally, the intercept and the weights of the variables included in the selected regression model were implemented in ArcGIS 10.1 using map algebra to generate categorical maps of fire severity. The spatial resolution of the maps was 25 m cell size, which was determined by the area of the sampled CBI plots. In seeking a useful map to assist land management purposes rather than a dichotomous response (presence or absence of fire impact), one of the methods adopted in literature is to divide the histogram of the probability map into different categories based on expert opinions [60]. This type of changing continuous data into two or more categories does not take into account the relative position of a case within the probability map and is neither fully automated nor statistically tested. Thus, in this study, we also considered alternative classification systems that use natural breaks, equal intervals and standard deviations. In addition, we also tested a manual classification of the probability values based on a direct relationship with the categories established by the CBI protocol: Unburned ($CBI = 0$), Low ($CBI \leq 1$), Moderate ($1 < CBI \leq 2$), Moderate-High ($2 < CBI \leq 2.5$) and High ($CBI > 2.5$). The selection of the final classification approach was based on its suitability to the information and the scale of investigation.

Finally, it should be noted that due to the lack of pre-fire information our model is only valid in areas that were covered by homogeneous pine stands before fire occurrence, so the final map only represent estimates of severity for those areas.

Table 4. Methodological steps performed to estimate the logistic regression model and its validation.

Step	Input	Test	Purpose
Prior analysis	CBI data and ALS variables	Spearman's coefficient	Assess the statistical correlation
	CBI data.	Shapiro–Wilk.	Goodness-of-fit test of normality.
	CBI data splitting into two samples: Low and High burn severity. Cut-off CBI value 1.5.	Kruskal–Wallis.	Test whether two samples have different means.
		Logistic regression.	Predict the outcome of CBI categories based on predictor LiDAR variables.
Multivariate statistical analysis	Training dataset (70%) of CBI data and ALS variables.	Log-likelihood (−2LL) and Chi X ² . Hosmer and Lemeshow	Assess the fitting of the logistic regression model to the observed data.
		Cox and Snell and Nagelkerke pseudo-R ² .	Estimate the strength of the relationship between the explanatory variables and the outcome.
		Wald and chi-square.	Assess the significance of each independent variable incorporated in the model and globally.
Validation analysis	Validation dataset (30%) of CBI data and ALS variables. Cut-off 0.6.	Cross tabulations, Cohen's Kappa index and ROC curves.	Analyze the performance of the binary classification.

3. Results

3.1. Correlation Analysis and Logistic Regression Model

Among the 51 explanatory variables derived from ALS data, 32 had a moderate–high significant Spearman's coefficient. Their correlation coefficients (ρ) are reported in Table 5. Height variables, such as kurtosis (0.788) or the 25th percentile (−0.767) had the strongest correlation with burn severity. In addition, the percentage of all returns above 1 m (−0.757) or the percentage of first returns above 1 m (−0.744) also presented high correlation coefficients. In the case of height metrics with negative coefficients, higher values of the variable indicate a higher amount of pulses returned by the canopy structure implying lower severity values. For example, a higher 25th percentile (e.g., 2 m) indicates that 75% of returns present height values above 2 m (lower severity) while a lower 25th percentile (e.g., 0.19 m) indicates that 75% of returns present height values above 0.19 m (higher severity). This was also the case for variables related to percentage of returns above a height threshold: a high percentage of returns from the total number of returns above 1 m implies low severity values while lower percentage of returns represent higher severity values. In contrast, variables as kurtosis and skewness presented positive coefficients indicating high number of returns from low heights (high severity) while negative coefficients corresponded to areas with a higher number of returns from higher heights (low severity).

Table 5. Spearman's coefficient (Rho) and Kruskal–Wallis (K.W.) chi-square values for selected variables with a statistical significance level p -value ≤ 0.01 .

ALS Variables	Rho	K.W. Chi ²	ALS Variables	Rho	K.W. Chi ²
Elev_kurtosis	0.788	54.169	Percentage first returns above 3.00	−0.690	39.927
Elev_P25	−0.767	64.776	Ratio_All_returns_3m	−0.690	39.797
Elev_P30	−0.764	63.550	Elev_mean	−0.684	34.138
%_All_returns_1m	−0.757	56.715	Elev_P60	−0.674	42.868
Elev_P20	−0.754	68.566	%_First_returns_mean	−0.673	62.590
Elev_P40	−0.752	63.802	Canopy relief ratio	−0.671	57.964
Elev_skewness	0.747	64.611	Ratio_All_returns_mean	−0.661	64.776
%_First_returns_1m	−0.744	52.460	Elev_P70	−0.653	33.988
Ratio_All_returns_1m	−0.742	51.915	Elev_P75	−0.649	29.533
%_All_returns_2m	−0.736	51.570	Elev_IQ	−0.637	29.544
%_Class_Unassigned	−0.729	54.131	Elev_P80	−0.631	24.535
%_Class_Ground	0.729	54.131	%_All_returns_mean	−0.630	67.337
Elev_P50	−0.728	57.146	%_num_of_ret 1	0.623	30.906
%_First_returns_2m	−0.723	47.116	%_num_of_ret 2	−0.622	31.020
Ratio_All_returns_2m	−0.722	46.904	Elev_AAD	−0.614	16.834
Percentage all returns above 3.00	−0.702	43.290	Elev_P90	−0.608	16.450

Concerning the differences between the two groups (null or low severity and high severity), the 20th percentile (68.566) and the percentage of all returns above mean (67.337) presented the highest Kruskal–Wallis chi-square values, followed by the 25th percentile and the percentage of all returns above mean divided by all first returns (64.776 in both cases). In general, low height percentiles (between 40th and 20th), height distribution measures such as skewness, and the percentage of returns above a certain threshold showed statistically significant differences between high and low severity values (Table 6). According to these results, differences between severities are identified by ALS variables that are commonly assumed to be related to vegetation structure.

The logistic regression models created with the training dataset using a forward stepwise method to select the explanatory variables are summarized in Table 7. The $-2LL$ and the likelihood ratio (LR) test with chi-square (X^2) distribution evaluate the fitting of the logistic regression model to the observed data. Smaller $-2LL$ values and higher X^2 values indicate a better fitting with a statistical significance ($p \leq 0.05$) [52,54]. Similar to linear regression, the pseudo- R^2 values show approximately how much variation is explained by the model.

The final model selected was the second (step 2) due to the principle of parsimony, and because the rest of the models included no significant variables. Nagelkerke's R^2 suggests that this model explains roughly 66% of the variation and is composed by the percentage of all returns above 1 m and the canopy relief ratio as significant ALS-derived variables. The Hosmer and Lemeshow's goodness-of-fit test confirms, since chi-square is not significant, that model two is an adequate fit. Table 8 shows the statistical significance of the individual predictors that entered in the selected model. The β coefficients of both variables are significant and negative, thereby indicating that the decrease in canopy relief ratio values and in the percentage of all returns above 1 m is associated with the increase in fire severity. The canopy relief ratio is a quantitative descriptor of the relative shape of the canopy

(Table 3) ranging between 0 and 1 and reflecting the degree to which canopy surfaces are in the upper (>0.5) or the lower (<0.5) portion of the height interval [61]: for high burn severity, the canopy relief ratio value tends to 0, whereas the value is close to 0.5 for low burn severity (see Figure 2). The percentage of all returns above 1 m is sensitive to the amount of biomass located above the understory layer. If the percentage is low it means that treetops have been scorched by fire, with only tree trunks and some branches remaining, which indicates high burn severities.

Table 6. Most important variables related to the selected CBI field plots used as example in Figures 4 and 5.

ALS Variables	CBI Plots					
	(a) Unburned (CBI = 0.0)	(b) Low (CBI = 1.0)	(c) Moderate (CBI = 1.5)	(d) Moderate (CBI = 2.0)	(e) Moderate–High (CBI = 2.5)	(f) High (CBI = 3.0)
Elev_skewness	−0.25	0.01	0.69	1.21	3.27	4.68
Elev_kurtosis	1.49	1.25	1.86	2.59	12.13	23.91
Elev_P20	0.19	0.08	0.03	0.01	0.01	0.00
Elev_P25	2.38	0.20	0.05	0.02	0.02	0.01
Elev_P30	4.05	1.80	0.07	0.03	0.02	0.01
Elev_P40	5.45	5.70	1.08	0.06	0.04	0.02
Elev_P50	6.37	7.11	3.38	0.09	0.05	0.03
Canopy relief ratio	0.44	0.42	0.29	0.20	0.08	0.06
%_All_returns_1m	68.77	59.15	46.43	25.76	8.14	4.36
Ratio_All_returns_mean	70.50	65.80	50.48	30.14	9.63	4.61
%_First_returns_mean	70.50	65.60	50.48	30.14	9.63	4.61
%_All_returns_mean	57.57	52.65	40.56	25.76	8.14	4.36

Table 7. Logistic regression models including the log likelihood ($-2LL$), model χ^2 test (LR), degrees of freedom ($d.f.$), significance level (P), pseudo- R^2 statistics and Hosmer and Lemeshow test.

Step/model	$-2LL$	Model χ^2 Test (LR)			Pseudo- R^2 Nagelkerke's R^2	Hosmer and Lemeshow		
		χ^2	d.f.	$P(>\chi^2)$		χ^2	d.f.	P
1	91.871	56.429	1	0.000	0.531	31.607	8	0.000
2	73.290	75.011	2	0.000	0.658	3.663	8	0.886
3	53.845	94.456	3	0.000	0.770	3.519	8	0.898
4	42.790	105.511	4	0.000	0.826	1.139	8	0.997
5	35.507	112.794	5	0.000	0.860	0.369	8	1.000
6	29.076	119.225	6	0.000	0.889	3.088	8	0.929
7	29.681	118.620	5	0.000	0.886	3.704	8	0.883
8	21.329	126.972	6	0.000	0.921	1.552	8	0.992
9	17.315	130.986	7	0.000	0.937	1.790	7	0.971
10	10.239	138.061	8	0.000	0.964	0.185	6	1.000
11	10.995	137.306	7	0.000	0.961	0.250	6	1.000

Table 8. β coefficients, Wald test values, degrees of freedom (*d.f.*) and their significance $p \leq 0.05$ computed for the variables of the selected regression model.

Independent Variables	β	Standard Error	Wald Test	<i>d.f.</i>	Signif.
Canopy relief ratio	−12.236	3.451	12.571	1	0.000
Percentage all returns above 1.00	−0.055	0.013	17.620	1	0.000
Constant	6.925	1.566	19.546	1	0.000

3.2. Model Validation

The discrimination ability of the logistic regression model was tested by a cross-tabulation between observed and predicted high (1) and low (0) severity cases. Table 9 shows the percentage of agreement between observed and predicted values for both training and validation samples. Value 1 was assigned to the predicted value when the obtained probability was greater than the cutoff value (*i.e.*, 0.6) quantified as the proportion of CBI plots classified as 1. Similar accuracies are achieved for both training and validation datasets with almost 85% of the cases being correctly classified. Furthermore, Cohen's Kappa index summarized in Table 10 demonstrates a good agreement for both training and validation datasets (0.681 and 0.565, respectively).

Table 9. Observed and predicted cross-tabulation for both training and validation datasets.

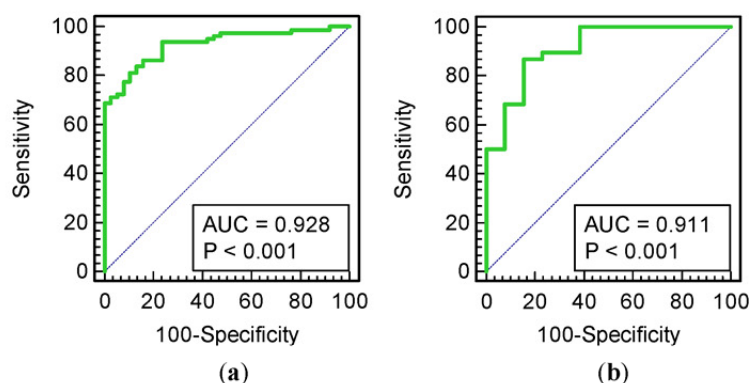
Training Dataset						Validation Dataset					
		Predicted						Predicted			
		Low	High	Sum	% Correct			Low	High	Sum	% Correct
Observed	Low	32	6	38	84.2	Observed	Low	8	5	13	61.9
	High	11	69	80	86.3		High	3	35	38	92.1
	Sum	43	75	118	85.6		Sum	11	40	51	84.3

Table 10. Cohen's Kappa index (K) for both training and validation datasets.

	K Training	K Validation
Value	0.681	0.565
Standad error	0.071	0.136
Signif.	0.000	0.000
Number of cases	118	51

Figure 6 presents the ROC curves and AUC values for both, the training and validation subsets of CBI plots. ROC curves are fairly similar and, consequently, very small differences of AUC values are observed with an asymptotic significance less than 0.05. Since, both the classification matrix and AUC values indicated that the modelling procedure carried out at sample plot scales has not suffered from over-fitting and the model demonstrates robustness.

Figure 6. ROC curves and AUC values of the regression model. Fitted ROC curve in green. Dashed line for an uninformative test (sensitivity + specificity = 1). (a) ROC Training Dataset; (b) ROC Validation Dataset.



3.3. Fire Severity Mapping

The final result of the regression process is the implementation of the model in a GIS environment with the aim of obtaining a spatial distribution of the fire severity probability (continuous values between 0 and 1) following Equation (2):

$$[P = \exp(6.925 - 12.236 \times \text{Canopy relief ratio} - 0.055 \times \text{Percentage of all returns above 1 m}) / 1 + \exp(6.925 - 12.236 \times \text{Canopy relief ratio} - 0.055 \times \text{Percentage of all returns above 1 m})] \quad (2)$$

The more these numbers are close to 1, the higher the likelihood of finding a zone impacted by high fire severity. As mentioned above, a threshold of 0.6 (proportion of 1 in the sample) can be used to distinguish between the two categories used to generate the model: areas where fire severity is null or low or areas with high fire severity. However, for management purposes, it is sometimes more useful to develop categorical maps. A few trials showed that natural breaks-based, equal intervals and standard deviation classification systems based on the probability histogram values resulted in misleading maps. Since Spearman's coefficient between CBI values and P probability showed a high correlation (0.807, p -value ≤ 0.01) the well-established CBI categories were used to divide the probability map in unburned, low, moderate, moderate-high and high fire severity: the average p values for each of the CBI classes were computed and subsequently used as central value for each severity range. This helped ensuring a better match between the observed values and those predicted by the logistic model. Table 11 presents the established categories of CBI value and the probability values assigned to them.

According to the above classification, Table 12 shows the percentage of fire severity within the fire perimeter by fire location. The results indicate that in Zuera the percentage of area affected by high fire severity was higher (91.16% of the total area inside of the fire perimeter). Conversely, Aliaga presented the highest percentage of area unaffected by fire, approximately 5.01%.

Table 11. The classification system used to produce fire severity categories and the proportions of plots by category.

CBI Value	Class Name	Probability Range	% of Plots Categorized According to the CBI Method	% of Plots Categorized According to the New Probability Ranges
CBI = 0	Unburned	0.00–0.30	24.85	21.89
CBI ≤ 1	Low	0.31–0.40	1.18	2.96
1 < CBI ≤ 2	Moderate	0.41–0.60	15.98	8.88
2 < CBI ≤ 2.5	Moderate–High	0.61–0.80	13.61	11.83
CBI > 2.5	High	0.81–1.00	44.38	54.44

Table 12. Percentage of area by fire severity classification and fire.

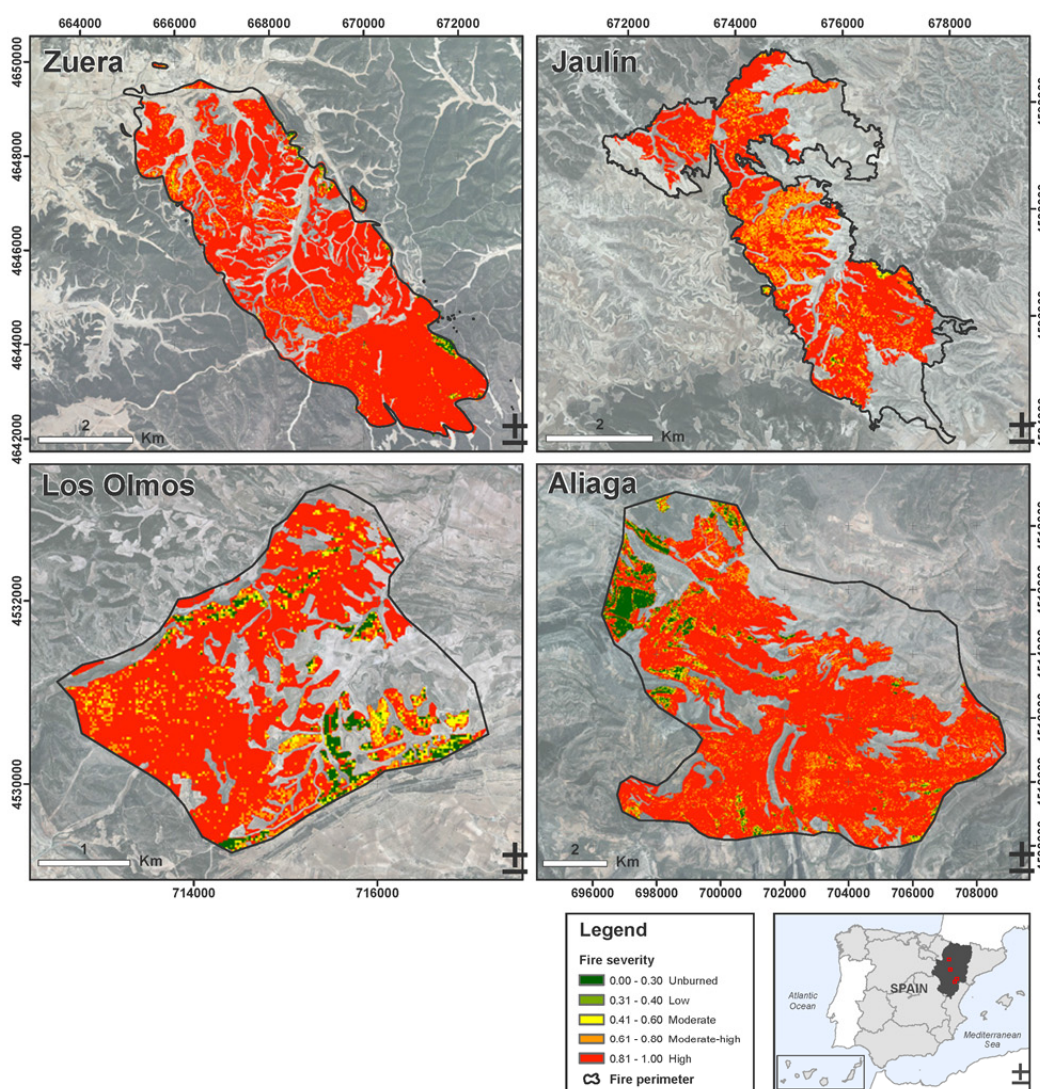
Class Name/Fire	Zuera	Jaulín	Aliaga	Los Olmos
Unburned	1.03%	0.30%	5.01%	4.67%
Low	0.41%	0.39%	1.31%	1.77%
Moderate	1.21%	3.56%	3.58%	5.04%
Moderate–High	6.19%	19.76%	10.04%	10.99%
High	91.16%	76.00%	80.06%	77.53%

Figure 7 presents the probability map generated in a raster-GIS environment. Most of the area was affected by high or moderate–high burn severity, whereas low burn severity areas were mainly encountered along the fire borders. As mentioned above, our model fits well only in areas that were covered by homogeneous and continuous pine stands previous to fire occurrence and is only applicable to these areas. The lack of pre-fire information resulted in assigning high severity values to sparse canopy woodlands and crop lands which had to be masked out for the analysis.

4. Discussion

In the Mediterranean basin fire is a natural and historical element. However, in recent decades, fire recurrence and magnitude are profoundly altering forest ecosystems, with forest degradation being the most immediate effect. The main objective of this study was to evaluate the suitability of ALS-derived variables to estimate fire severity in four burned areas in Aragón, Spain. Our aim was to establish relationships between CBI data collected in the field and different variables derived from ALS point clouds. However, due to the absence of pre-fire ALS flights the two data sources provide information of different characteristics. On one hand, the CBI index estimates the change in the ecosystem after fire with respect to a previous situation using qualitative parameters that are then encoded into ranges of burn severity. On the other hand, the ALS data only provided estimations of vegetation structure after fire and not the vegetation change. Thus, we assumed homogeneous and continuous pre-fire vegetation in all field plots as reported by Tanase *et al.* [20]. Given these assumptions, a strong relationship was found between ALS-derived variables related to vegetation structural parameters and fire severity values, allowing the establishment of a model for fire severity mapping.

Figure 7. Fire severity maps obtained with the logistic method for pine cover in the four study locations. High spatial resolution orthophotography (2012) is used as backdrop.



Kane *et al.* [34] noted the significant effects of fire on forest structure by using ALS data with Landsat-derived estimates of fire severity to measure the impact of fires over large areas. As fire severity increased, the canopy cover decreased while the number of tree clumps increased, indicating progressive canopy fragmentation. The work confirmed the utility of Landsat-based fire severity estimates and the utility of high resolution of ALS data to measure the structural change resulting from that process. This is one of the first approaches to fire severity assessment using ALS data; however, according to our knowledge, our work represents for the first time fire severity modelled using ALS-derived variables in relation to CBI field data, joining a small but growing body of literature focused on deriving fire severity estimates from ALS data [21,24,34,51].

The lack of pre-fire forest structure information derived from ALS data may be solved in the future by periodical acquisitions within the PNOA mission. Wang *et al.* [62] demonstrated high fire-severity

accuracies (84%) in a rangeland ecosystem using pre- and post-fire ALS data. Their approach was based on evaluating changes in vegetation average height using pre- and post-fire ALS data, which was subsequently related to biomass combustion. However, the growing availability of ALS datasets with increasing nominal point density, combined with methods such as the one discussed in this paper, can assist forest management planning even when pre-fire ALS data are not available.

The way fire severity is distributed is a key factor in quantifying the impact of fires and post-fire management [9], as well as the ecosystem responses. The severity mapping attempts to identify problematic areas and help managers in allocating resources and restoration efforts and thus decreasing time and economic costs associated with field-work. Logistic regression was selected due to its ability to work with different types of independent variables, even when they are auto-correlated as well as due to its independency from data statistical distribution [55]. The method allowed for generating severity maps with up to 85.8% prediction accuracy when assessing spatial distribution of fire impact within pine forests. In addition, this research contributes to the creation of sensible theories about which explanatory variables are most important for fire severity estimation from low density ALS data, such as ALS measures of canopy shape and percentage of all returns above one meter.

Our research, like many ALS-based studies, lacked concurrent field and ALS acquisitions [24]. However, the evergreen character of species present in the study area compensated for such discrepancies. On the contrary, at the time of the ALS flight, some of the affected forests were felled and removed by forest management agencies. Field plots located in such areas were removed from analysis, considerably reducing the initial sampling to only 169 plots. In addition, the time gap could have diminished the accuracy of burn severity predictions due to the increasing vegetation recovery, especially in sprouting shrubs such as *Quercus coccifera* L.

In order to determine whether ALS-derived fire severity estimations produce a better relationship than differenced Normalized Burn Ratio (dNBR) and its relative form (RdNBR) Landsat TM-based indices, the results of applying these indices, computed by Tanase *et al.* [20], and the results of our logistic model were correlated with CBI field plots data. The dNBR (Figure 8a) resulted in an R^2 of 0.67, the RdNBR model (Figure 8c) achieved a R^2 of 0.68 and the logistic model resulted in an R^2 of 0.63. These coefficients of determination indicate that the relationship between observed data (CBI values) and the predicted severity from ALS logistic model is good but it does not exceed the coefficients obtained for the RdNBR and dNBR. Nevertheless, the ALS-based model, lacking pre-fire information, yields similar accuracy to optical approaches that used pre- and post-fire information. On the other hand, the relationship between our logistic model and Landsat-derived indices is moderate (Figure 8b,d), being better with RdNBR ($R^2 = 0.58$) than with dNBR. Although all the approaches obtained good results, clearly there is still no perfect index for mapping fire effects. We therefore need to further investigate in this direction.

Accurately mapping the distribution of severity patches is important for site level recovery projects and for understanding overall landscape patterns created by fire. Thus, a detailed examination of the relation of established CBI classes and the values obtained with the logistic model, RdNBR and dNBR were performed. Then, we created and mapped five severity categories based on the mean values obtained at each CBI class.

Figure 8. Scatterplots depicting the relationship between three remotely sensed severity metrics (RdNBR, dNBR, and logistic model) and CBI. The black lines represent the regression trend. The coefficient of determination (R^2) and the equation of the line are shown for each fit. (a) CBI vs. dNBR; (b) Logistic model vs. dNBR; (c) CBI vs. RdNBR; (d) Logistic model vs. RdNBR; (e) CBI vs. Logistic model.

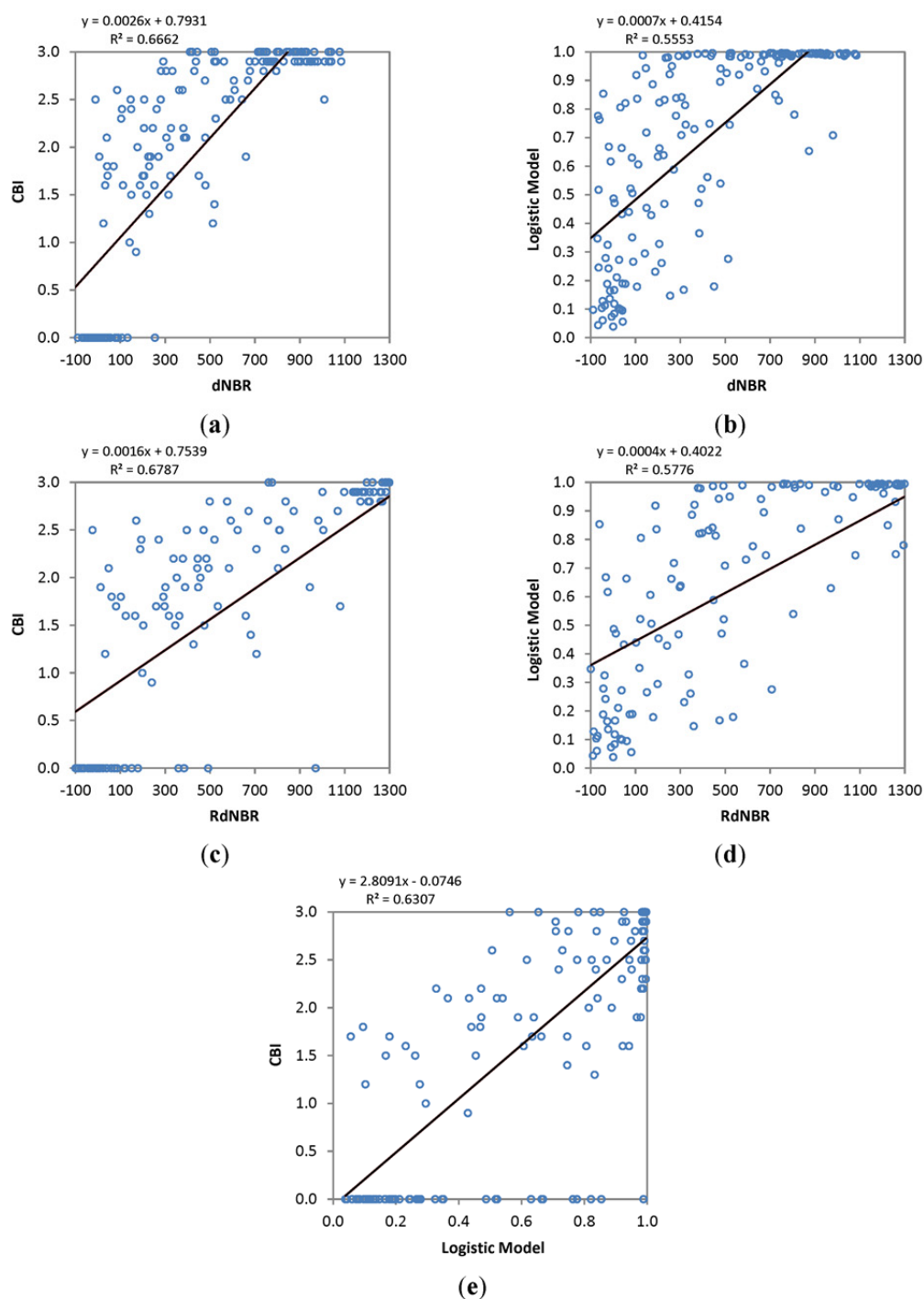


Table 13 shows the category ranges obtained, the Kappa index of agreement between the CBI classes and the categories obtained with the three different approaches and the percentage of agreement in each category. As can be observed, the logistic regression approach led to under-representing low to moderate–high severity patches but presents better accuracy in the high severity class (89.33% accuracy) than RdNBR (86.67%) and dNBR (77.33%).

Table 13. Ranges for the three remotely sensed fire severity metrics according to CBI values, accuracies (%) and Kappa index for each class (p -value ≤ 0.01).

CBI Value	Class Name	Fire Severity Metrics			Accuracies (%)		
		Logistic Model Probability	RdNBR	dNBR	Logistic Model Probability (Kappa = 0.42)	RdNBR (Kappa = 0.60)	dNBR (Kappa = 0.55)
CBI = 0	Unburned	<0.30	<118	<81	66.67	83.33	85.71
CBI ≤ 1	Low	0.30	297	198	0.00	100.00	100.00
1 < CBI ≤ 2	Moderate	0.40	451	291	18.52	29.63	33.33
2 < CBI ≤ 2.5	Moderate–high	0.60	894	545	13.04	47.83	34.78
CBI > 2.5	High	>0.80	>894	>545	89.33	86.67	77.33
Overall accuracy					60.95	71.60	66.86

In our opinion, minimizing classification errors in high severity class is beneficial to land managers as it allows a better identification of areas that are severely burned [63]. However, to improve classification results, other methods such as random forests and nearest neighbor imputation could allow multiple severity classes mapping.

The integration of optical remotely sensed imagery, well suited for capturing horizontally distributed forest conditions, along with ALS variables that are more appropriate for capturing vertical forest structure, may improve fire severity assessment [64]. Finally, we are of the opinion that the generation of indices derived from pre- and post-fire ALS variables will considerably improve fire severity mapping.

5. Conclusions

This study presents a new methodological approach combining ALS-derived variables and field-assessed CBI, in order to estimate the impact of wildland fire across four sites in Aragón, northeastern Spain. Our work is unique in examining fire severity within this particular setting, *i.e.*, Mediterranean pine forests and low-density ALS data (0.5 pulses/m²). We have demonstrated that ALS-derived variables from plot-level distributions of pulse return heights provide initial support for assessing the impact caused by fire. Height variables such as kurtosis or the 25th percentile of returns heights presented the highest correlation with CBI values (0.788 and, respectively, -0.767). The difference between high and low severity classes was best estimated using the 20th percentile and the percentage of all returns above mean which presented the highest Kruskal–Wallis chi-square values (68.566 and 67.337, respectively). Low percentiles (between 40th and 20th), skewness, and some percentages of ALS returns above a threshold, such as the average height, also showed significant differences between the two groups of severity.

The relationships between CBI data and a set of ALS point cloud variables have been assessed by means of forward stepwise logistic regression. The prediction power of the obtained model was tested

using independent validation samples. The results show a fairly acceptable overall accuracy, confirming that logistic regression is an effective tool for fire severity analysis. Continuous fire severity maps were derived based on the probabilities obtained from the logistic model at plot level. Next, the average of those probabilities for each severity class allowed for the split of the map in fire severity classes. The lack of pre-fire forest structure information can be a handicap that could be solved by periodical acquisition of ALS datasets in the coming years.

Acknowledgments

This work has been financed by the Government of Aragón, Department of Science, Technology and University (FPI Grant BOA 30, 11/02/2011). The ALS data were provided by the National Center for Geographic Information of Spain. Francisco Palú, Marco Lorenzo and Emilio Pérez Aguilar from the Provincial Environment Service, Government of Aragón are also acknowledged for their help with the forest management activities carried out after the fires.

Author Contributions

Juan de la Riva had the original idea for the study. Antonio L. Montealegre and M.T. Lamelas developed the methodology and performed the analysis. Field data was provided by Mihai A. Tanase. Antonio L. Montealegre wrote the manuscript, incorporating suggestions from all authors, who approved the final manuscript.

Conflicts of Interest

The authors declare no conflict of interest.

References

1. Bond, W.J.; Keeley, J.E. Fire as a global “herbivore”: The ecology and evolution of flammable ecosystems. *Trends Ecol. Evol.* **2005**, *20*, 387–394.
2. Amraoui, M.; Liberato, M.L.R.; Calado, T.J.; DaCamara, C.C.; Coelho, L.P.; Trigo, R.M.; Gouveia, C.M. Fire activity over Mediterranean Europe based on information from Meteosat-8. *For. Ecol. Manag.* **2013**, *294*, 62–75.
3. Ichoku, C.; Giglio, L.; Wooster, M.J.; Remer, L.A. Global characterization of biomass-burning patterns using satellite measurements of fire radiative energy. *Remote Sens. Environ.* **2008**, *112*, 2950–2962.
4. Oliveira, S.; Oehler, F.; San-Miguel-Ayanz, J.; Camia, A.; Pereira, J.M.C. Modeling spatial patterns of fire occurrence in Mediterranean Europe using multiple regression and random forest. *For. Ecol. Manag.* **2012**, *275*, 117–129.
5. Pausas, J.G.; Llovet, J.; Rodrigo, A.; Vallejo, R. Are wildfires a disaster in the Mediterranean basin?—A review. *Int. J. Wildland Fire* **2008**, *17*, 713–723.
6. San-Miguel-Ayanz, J.; Moreno, J.M.; Camia, A. Analysis of large fires in European Mediterranean landscapes: Lessons learned and perspectives. *For. Ecol. Manag.* **2013**, *294*, 11–22.

7. Eleazar, M.J.; Enríquez, E.; Gallar, J.J.; Jemes, V.; López, M.; Mateo, M.L.; Muñoz, A.; Parra, P.J. *Los Incendios Forestales en España, Decenio 2001–2010*; Área de Defensa contra Incendios Forestales (ADCIF) del Ministerio de Agricultura, Alimentación y Medio Ambiente: Madrid, Spain, 2012; p. 138.
8. Collins, R.D.; de Neufville, R.; Claro, J.; Oliveira, T.; Pacheco, A.P. Forest fire management to avoid unintended consequences: A case study of Portugal using system dynamics. *J. Environ. Manag.* **2013**, *130*, 1–9.
9. Chuvieco, E. *Earth Observation of Wildland Fires in Mediterranean Ecosystems*; Springer: Alcalá de Henares, Spain, 2009; p. 257.
10. De Santis, A.; Chuvieco, E. Burn severity estimation from remotely sensed data: Performance of simulation vs. empirical models. *Remote Sens. Environ.* **2007**, *108*, 422–435.
11. Landscape Assessment (LA) Sampling and Analysis Methods. Available online: http://www.fs.fed.us/rm/pubs/rmrs_gtr164/rmrs_gtr164_13_land_assess.pdf (accessed on 5 May 2014).
12. Epting, J.; Verbyla, D.; Sorbel, B. Evaluation of remotely sensed indices for assessing burn severity in interior Alaska using Landsat TM and ETM+. *Remote Sens. Environ.* **2005**, *96*, 328–339.
13. Kasischke, E.S.; Turetsky, M.R.; Ottmar, R.D.; French, N.H.F.; Hoy, E.E.; Kane, E.S. Evaluation of the composite burn index for assessing fire severity in Alaskan black spruce forests. *Int. J. Wildland Fire* **2008**, *17*, 515–526.
14. De Santis, A.; Chuvieco, E. GeoCBI: A modified version of the composite burn index for the initial assessment of the short-term burn severity from remotely sensed data. *Remote Sens. Environ.* **2009**, *113*, 554–562.
15. Loboda, T.V.; French, N.H. F.; Hight-Harf, C.; Jenkins, L.; Miller, M.E. Mapping fire extent and burn severity in Alaskan tussock tundra: An analysis of the spectral response of tundra vegetation to wildland fire. *Remote Sens. Environ.* **2013**, *134*, 194–209.
16. Hall, R.J.; Freeburn, J.T.; de Groot, W.J.; Pritchard, J.M.; Lynham, T.J.; Landry, R. Remote sensing of burn severity: experience from western Canada boreal fires. *Int. J. Wildland Fire* **2008**, *17*, 476–489.
17. Veraverbeke, S.; Verstraeten, W.W.; Lhermitte, S.; Goossens, R. Evaluating landsat thematic mapper spectral indices for estimating burn severity of the 2007 Peloponnese wildfires in Greece. *Int. J. Wildland Fire* **2010**, *19*, 558–569.
18. Hudak, A.T.; Morgan, P.; Bobbitt, M.J.; Smith, A.M.S.; Lewis, S.A.; Lentile, L.B.; Robichaud, P.R.; Clark, J.T.; McKinley, R.A. The relationship of multispectral satellite imagery to immediate fire effects. *Fire Ecol.* **2007**, *3*, 64–90.
19. Van Wagendonk, J.W.; Root, R.R.; Key, C.H. Comparison of AVIRIS and Landsat ETM+ detection capabilities for burn severity. *Remote Sens. Environ.* **2004**, *92*, 397–408.
20. Tanase, M.; de la Riva, J.; Pérez-Cabello, F. Estimating burn severity at the regional level using optically based indices. *Can. J. For. Res.* **2011**, *41*, 863–872.
21. Bergen, K.M.; Goetz, S.J.; Dubayah, R.O.; Henebry, G.M.; Hunsaker, C.T.; Imhoff, M.L.; Nelson, R.F.; Parker, G.G.; Radeloff, V.C. Remote sensing of vegetation 3-D structure for biodiversity and habitat: Review and implications for lidar and radar spaceborne missions. *J. Geophys. Res.* **2009**, *114*, G00E06, doi:10.1029/2008JG000883.

22. Tanase, M.A.; Santoro, M.; Wegmüller, U.; de la Riva, J.; Pérez-Cabello, F. Properties of X-, C- and L-band repeat-pass interferometric SAR coherence in Mediterranean pine forests affected by fires. *Remote Sens. Environ.* **2010**, *114*, 2182–2194.
23. Tanase, M.; de la Riva, J.; Santoro, M.; Pérez-Cabello, F.; Kasischke, E. Sensitivity of SAR data to post-fire forest regrowth in Mediterranean and boreal forests. *Remote Sens. Environ.* **2011**, *115*, 2075–2085.
24. Kane, V.R.; Lutz, J.A.; Roberts, S.L.; Smith, D.F.; McGaughey, R.J.; Povak, N.A.; Brooks, M.L. Landscape-scale effects of fire severity on mixed-conifer and red fir forest structure in Yosemite National Park. *For. Ecol. Manag.* **2013**, *287*, 17–31.
25. Vosselmann, G.; Maas, H.G. *Airborne and Terrestrial Laser Scanning*; Whittles Publishing: Dunbeath, UK, 2010.
26. Lefsky, M.A.; Cohen, W.B.; Acker, S.A.; Parker, G.G.; Spies, T.A.; Harding, D. Lidar remote sensing of the canopy structure and biophysical properties of douglas-fir western hemlock forests. *Remote Sens. Environ.* **1999**, *70*, 339–361.
27. Naesset, E.; Økland, T. Estimating tree height and tree crown properties using airborne scanning laser in a boreal nature reserve. *Remote Sens. Environ.* **2002**, *79*, 105–115.
28. Andersen, H.E.; McGaughey, R.J.; Reutebuch, S.E. Estimating forest canopy fuel parameters using LIDAR data. *Remote Sens. Environ.* **2005**, *94*, 441–449.
29. Zolkos, S.G.; Goetz, S.J.; Dubayah, R. A meta-analysis of terrestrial aboveground biomass estimation using lidar remote sensing. *Remote Sens. Environ.* **2013**, *128*, 289–298.
30. Kane, V.R.; Gersonde, R.F.; Lutz, J.A.; McGaughey, R.J.; Bakker, J.D.; Franklin, J.F. Patch dynamics and the development of structural and spatial heterogeneity in Pacific Northwest forests. *Can. J. For. Res.* **2011**, *41*, 2276–2291.
31. Reutebuch, S.E.; Andersen, H.-E.; McGaughey, R.J. Light detection and ranging (LIDAR): An emerging tool for multiple resource inventory. *J. For.* **2005**, *103*, 286–292.
32. Hudak, A.T.; Evans, J.S.; Smith, A.M.S. LiDAR utility for natural resource managers. *Remote Sens.* **2009**, *1*, 934–951.
33. Asner, G.P.; Hughes, R.F.; Mascaró, J.; Uowolo, A.L.; Knapp, D.E.; Jacobson, J.; Kennedy-Bowdoin, T.; Clark, J.K. High-resolution carbon mapping on the million-hectare Island of Hawaii. *Front. Ecol. Environ.* **2011**, *9*, 434–439.
34. Kane, V.R.; North, M.P.; Lutz, J.A.; Churchill, D.J.; Roberts, S.L.; Smith, D.F.; McGaughey, R.J.; Kane, J.T.; Brooks, M.L. Assessing fire effects on forest spatial structure using a fusion of Landsat and airborne LiDAR data in Yosemite National Park. *Remote Sens. Environ.* **2013**, *2013*, doi:org/10.1016/j.rse.2013.07.041.
35. Agca, M.; Popescu, S.C.; Harper, C.W. Deriving forest canopy fuel parameters for loblolly pine forests in eastern Texas. *Can. J. For. Res.* **2011**, *41*, 1618–1625.
36. Erdody, T.L.; Moskal, L.M. Fusion of LiDAR and imagery for estimating forest canopy fuels. *Remote Sens. Environ.* **2010**, *114*, 725–737.
37. Riaño, D.; Chuvieco, E.; Condés, S.; González-Matesanz, J.; Ustin, S.L. Generation of crown bulk density for *Pinus sylvestris* L. from lidar. *Remote Sens. Environ.* **2004**, *92*, 345–352.

38. Riaño, D.; Chuvieco, E.; Ustin, S.L.; Salas, J.; Rodríguez-Pérez, J.R.; Ribeiro, L.M.; Viegas, D.X.; Moreno, J.M.; Fernández, H. Estimation of shrub height for fuel-type mapping combining airborne LiDAR and simultaneous color infrared ortho imaging. *Int. J. Wildland Fire* **2007**, *16*, 341–348.
39. Mutlu, M.; Popescu, S.C.; Zhao, K. Sensitivity analysis of fire behavior modeling with LiDAR-derived surface fuel maps. *For. Ecol. Manag.* **2008**, *256*, 289–294.
40. Skowronski, N.S.; Clark, K.L.; Duveneck, M.; Hom, J. Three-dimensional canopy fuel loading predicted using upward and downward sensing LiDAR systems. *Remote Sens. Environ.* **2011**, *115*, 703–714.
41. Vicente-Serrano, S.M.; Lasanta, T.; Gracia, C. Aridification determines changes in forest growth in *Pinus halepensis* forests under semiarid Mediterranean climate conditions. *Agric. For. Meteorol.* **2010**, *150*, 614–628.
42. Spanish National Plan for Aerial Orthophotography (PNOA). Available online: http://www.ign.es/PNOA/vuelo_lidar.html (accessed on 30 December 2013).
43. MCC-LiDAR Multiscale Curvature Classification for LiDAR data. Available online: <http://sourceforge.net/p/mcelidar/wiki/Home/> (accessed on 2 April 2014).
44. Evans, J.S.; Hudak, A.T. A Multiscale curvature algorithm for classifying discrete return LiDAR in forested environments. *IEEE Trans. Geosci. Remote Sens.* **2007**, *45*, 1029–1038.
45. Montealegre, A.; Lamelas, T.; de la Riva, J. Evaluación de Métodos de Filtrado Para la Clasificación de la Nube de Puntos del Vuelo LiDAR PNOA. In *Teledetección. Sistemas Operacionales de Observación de la Tierra*; Fernández-Renau González-Anleo, A., de Miguel Llanes, E., Eds.; INTA: Madrid, Spain, 2013; pp. 184–187.
46. Hutchinson, M.F.; Xu, T.; Stein, J.A. Recent Progress in the ANUDEM Elevation Gridding Procedure. In *Proceedings of the Geomorphometry, 2011*, Redlands, CA, USA, 7–9 September 2011; pp. 19–22.
47. Fusing LIDAR Data, Photographs, and Other Data Using 2D and 3D Visualization Techniques. Available online: http://www.fs.fed.us/pnw/olympia/silv/publications/opt/488_McGaugheyCarson2003.pdf (accessed on 29 April 2014).
48. Evans, J.; Hudak, A.; Faux, R.; Smith, A.M. Discrete return Lidar in natural resources: Recommendations for project planning, data processing, and deliverables. *Remote Sens.* **2009**, *1*, 776–794.
49. González-Olabarria, J.-R.; Rodríguez, F.; Fernández-Landa, A.; Mola-Yudego, B. Mapping fire risk in the model forest of urbión (Spain) based on airborne LiDAR measurements. *For. Ecol. Manag.* **2012**, *282*, 149–156.
50. Álvarez Cáceres, R. *Estadística Multivariante y No Paramétrica con SPSS: Aplicación a las Ciencias de la Salud*; Díaz de Santos: Madrid, Spain, 1994; p. 408.
51. Angelo, J.J.; Duncan, B.W.; Weishampel, J.F. Using lidar-derived vegetation profiles to predict time since fire in an oak scrub landscape in East-Central Florida. *Remote Sens.* **2010**, *2*, 514–525.
52. Hosmer, D.W.; Lemeshow, S. *Applied Logistic Regression*; Wiley: Hoboken, USA, 2000.
53. Hair, J.F.; Anderson, R.E.; Tatham, R.L.; Black, W.C. *Análisis Multivariante*, 5th ed.; Prentice Hall Iberia: Madrid, Spain, 1999.

54. Conoscenti, C.; Angileri, S.; Cappadonia, C.; Rotigliano, E.; Agnesi, V.; Märker, M. Gully erosion susceptibility assessment by means of GIS-based logistic regression: A case of Sicily (Italy). *Geomorphology* **2014**, *204*, 399–411.
55. Menard, S. *Logistic Regression: From Introductory to Advanced Concepts and Applications*; SAGE: Thousand Oaks, CA, USA, 2010; p. 392.
56. Lamelas, M.T.; Marinoni, O.; Hoppe, A.; de la Riva, J. Dollines probability map using logistic regression and GIS technology in the central Ebro Basin (Spain). *Environ. Geol.* **2008**, *54*, 963–977.
57. Beguería, S.; Lorente, A. *Landslide Hazard Mapping by Multivariate Statistics: Comparison of Methods and Case Study in the Pyrenees*; Instituto Pirenaico de Ecología, Contract No. EVG1-CT-1999–00007. Available online: http://damocles.irpi.pg.cnr.it/docs/reports/df_modelling.pdf (accessed on 29 April 2014).
58. Landis, J.R.; Koch, G.G. The measurement of observer agreement for categorical data. *Biometrics* **1977**, *33*, 159–174.
59. Hanley, J.A.; McNeil, B. The meaning and use of the área under a receiver operating characteristic (ROC) curve. *Radiology* **1982**, *143*, 29–36.
60. Ayalew, L.; Yamagishi, H. The application of GIS-based logistic regression for landslide susceptibility mapping in the Kakuda-Yahiko Mountains, Central Japan. *Geomorphology* **2005**, *65*, 15–31.
61. Parker, G.G.; Russ, M.E. The canopy surface and stand development: Assessing forest canopy structure and complexity with near-surface altimetry. *For. Ecol. Manag.* **2004**, *189*, 307–315.
62. Wang, C.; Glenn, N.F. Estimation of fire severity using pre- and post-fire LiDAR data in sagebrush steppe rangelands. *Int. J. Wildland Fire* **2009**, *18*, 848–856.
63. Miller, J.D.; Thode, A.E. Quantifying burn severity in a heterogeneous landscape with a relative version of the delta Normalized Burn Ratio (dNBR). *Remote Sens. Environ.* **2007**, *109*, 66–80.
64. Wulder, M.A.; White, J.C.; Alvarez, F.; Han, T.; Rogan, J.; Hawkes, B. Characterizing boreal forest wildfire with multi-temporal Landsat and LIDAR data. *Remote Sens. Environ.* **2009**, *113*, 1540–1555.

© 2014 by the authors; licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution license (<http://creativecommons.org/licenses/by/3.0/>).

3.6 Cartografía de tipos de combustible forestal

ABSTRACT

Fuel types mapping is a key factor in the fire ignition and propagation potential and constitutes a primary component of fire risk and forest management. In this sense, the activity in military training areas and shooting ranges implies an important human factor on forest fire occurrence. In this context, the objective of this research is Prometheus fuel types mapping combining multispectral optic remote sensing, LiDAR, SAR and in-situ data collected over 108 plots. In order to select the independent variables to be included in the digital classification process, a Kruskal-Wallis test was performed. The highest accuracy, 72.7% of global agreement in fuel type identification using a supervised classification and Kappa index of 0.7, was obtained combining optic and LiDAR data.

Cartografía de modelos de combustible mediante combinación de imágenes LiDAR, SAR y ópticas en el Centro de Adiestramiento “San Gregorio”

Montealegre Gracia, Antonio Luis ^{1,3,*}, Lamelas Gracia, María Teresa ^{2,3}, García-Martín, Alberto ^{2,3}, de la Riva Fernández, Juan ^{1,3} y Escribano Bernal, Francisco ²

¹ Departamento de Geografía y O.T., Universidad de Zaragoza, Pedro Cerbuna 12, 50009, Zaragoza, España; E-Mail: delariva@unizar.es (J.)

² Centro Universitario de la Defensa de Zaragoza, Academia General Militar, Ctra. de Huesca s/n, 50090, Zaragoza, España; E-Mail: tlamelas@unizar.es (M.T.), algarcia@unizar.es (A.), fescrib@unizar.es (F.)

³ GEOFOREST-IUCA, Universidad de Zaragoza, Pedro Cerbuna 12, 50009, Zaragoza, España.

* Autor Principal y responsable del trabajo; E-Mail: monteale@unizar.es (A.L.)

Abstract: La cartografía de los tipos de combustible forestal de un territorio es fundamental para la estimación tanto de la ignición como de las condiciones de propagación del fuego en caso de incendio, con lo que es indispensable para la gestión forestal, cobrando aún mayor importancia en los campos de maniobras militares, dado que las actividades desarrolladas en éstos suponen un incremento del riesgo de incendio. En este contexto, el objetivo de esta investigación ha sido el desarrollo de una metodología operativa para obtener una cartografía detallada de tipos de combustible basada en el modelo *Prometheus*, mediante la integración de información de teledetección óptica multiespectral, LiDAR, SAR y datos de 108 parcelas de campo usados como “verdad-terreno”. La selección de las variables independientes susceptibles de ser utilizadas en el proceso de clasificación digital se ha realizado mediante la prueba estadística de contraste de *Kruskal-Wallis*. La clasificación digital supervisada del área de estudio llevada a cabo, revela que la combinación de fuentes de información óptica y LiDAR produce el mejor resultado, con un 72,7% de acierto global en la identificación de las tipologías de combustibles y un índice *kappa* de 0,7.

Keywords: Modelo *Prometheus*; tipos de combustible; teledetección; imagen multiespectral; LiDAR; SAR; Clasificación digital

III Congreso Nacional de I+D en Defensa y Seguridad, 2015

1. Introducción

España sufre cada año las consecuencias de los incendios forestales, a pesar de que se ha observado una reducción en el número durante la década 2001-2010. Sin embargo, la ocurrencia de grandes incendios, superiores a 100 ha, ha aumentado debido a tres causas fundamentales [1]: i) el incremento de la combustibilidad como consecuencia de la transformación de los sistemas productivos y de los usos de suelo; ii) las condiciones meteorológicas extremas, con periodos prolongados de sequía, relacionadas con el cambio climático; y iii) la actividad humana que tiene lugar en el territorio, que cada vez es más intensa e implica en numerosas ocasiones un alto riesgo de incendio.

Más concretamente, los campos de tiro y maniobras del Ministerio de Defensa tienen un factor adicional en la ocurrencia de incendios forestales, ya que en ellos convergen una serie de condicionantes que entrañan un riesgo directo de ignición, por un lado ligado a las actividades desarrolladas, como son las prácticas de tiro con fuego real y la caída de proyectiles, entre otras, y por otro a la presencia de amplias superficies forestales dentro y en su entorno próximo. Por todo ello, es necesario disponer de un Subsistema de Prevención y Extinción de Incendios Forestales (SPEIF) en el ámbito de los campos de maniobra o propiedades similares del Ministerio de Defensa, tal y como se recoge en la Directiva 42/2010, de 12 de julio, de la Ministra de Defensa [2] como desarrollo de la disposición adicional octava de la Ley 43/2003, de 21 de noviembre, de Montes, modificada por la Ley 21/2015, de 20 de julio de 2015 [3]. Esto es fundamental para velar por la integridad del patrimonio forestal y natural en general, por las infraestructuras propias de este tipo de instalaciones y por las personas que desempeñan sus actividades en ese entorno.

En este sentido, la evaluación del riesgo de incendio, que es clave para la prevención y la planificación previa a la extinción, integra la estimación de: i) la ignición (humedad del combustible, factores humanos y naturales); ii) la propagación (propiedades físicas del combustible y condiciones topográficas y meteorológicas); y iii) la vulnerabilidad (socioeconómica y ecológica). Sin embargo, la gran diversidad de características que influyen en la combustibilidad de la vegetación obliga a desarrollar tipologías de combustibles caracterizados por una serie de parámetros cualitativos y/o cuantitativos que determinan unas condiciones concretas para la propagación del fuego [4]. En la actualidad la mayoría de los países ha desarrollado clasificaciones basadas en las ecuaciones de propagación de Rothermel [5], como el sistema de Estados Unidos (*National Fire Danger Rating System*, NFDRS), pero, en el entorno mediterráneo, una clasificación de especial interés por su adaptación a las condiciones regionales es la propuesta por el proyecto europeo *Prometheus* [6]. Este modelo presenta un marcado carácter estructural a la hora de modelizar el comportamiento del fuego, ya que tiene en cuenta la altura y la densidad de los elementos propagadores [4].

Las imágenes ópticas multiespectrales han sido habitualmente utilizadas en la cartografía de tipos de combustible a escala local y regional como un procedimiento alternativo frente a los costosos métodos de identificación a partir de fotografías aéreas o directamente en campo, aunque los sistemas de teledetección pasiva presentan ciertas limitaciones, dada su incapacidad para detectar el combustible localizado bajo el dosel vegetal. Además, la reflectividad que es captada por estos sensores no puede relacionarse con variables dasométricas como la altura de la vegetación, que es un parámetro imprescindible en la clasificación de tipos de combustible [7]. Este inconveniente puede ser solucionado con el uso de sensores activos LiDAR (*Light Detection And Ranging*) y SAR (*Synthetic*

III Congreso Nacional de I+D en Defensa y Seguridad, 2015

Aperture Radar), que proporcionan información de la estructura vertical de la vegetación. Los estudios orientados a la cartografía de combustibles basados en técnicas de clasificación digital de imágenes han demostrado cómo la combinación de información proveniente de sensores pasivos y activos mejora los resultados que se obtienen con cada uno de estos tipos de datos separadamente [8,9].

En este contexto, el objetivo general de este trabajo es establecer una metodología operativa que permita identificar los tipos de combustible según el modelo *Prometheus* a partir de información óptica multispectral, LiDAR y SAR, junto con trabajo de campo como apoyo al plan de prevención de incendios forestales del Centro de Adiestramiento “San Gregorio”. Los objetivos específicos de este trabajo son: (1) determinar cuál es la mejor combinación de fuentes de información para identificar los combustibles; y (2) producir una cartografía digital de combustibles de alta resolución.

2. Materiales y Métodos

2.1. Zona de estudio

La zona de estudio corresponde al Centro de Adiestramiento (CENAD) “San Gregorio” (Figura 1), que ocupa 33.839 ha en el sector central de la Depresión del Ebro (provincia de Zaragoza). Este territorio tiene un clima mediterráneo con rasgos continentales y ambiente semi-árido y unos suelos yesíferos pobres. Dominan los terrenos cubiertos por pastos de gramíneas y arbustos de hoja perenne de *Quercus coccifera* L., *Juniperus oxycedrus* L. subsp. *macrocarpa* (Sibth. y Sm.) Ball., y *Rosmarinus officinalis* L. Existen también masas arboladas de *Pinus halepensis* Mill. en el extremo norte y en rodales discontinuos de diverso tamaño siguiendo el límite este-sureste.

2.2. Datos de teledetección y trabajo de campo

Con respecto a los datos de teledetección, se manejaron tres tipos de informaciones (Tabla 1): i) una imagen multispectral del satélite SPOT-5 proporcionada por el Plan Nacional de Teledetección (PNT); ii) las nubes de puntos LiDAR del Plan Nacional de Ortofotografía Aérea (PNOA); y iii) tres imágenes SAR captadas por los satélites ENVISAT-1 y ERS-2 de la Agencia Espacial Europea (ESA).

Tipo de datos	Plataforma	Fecha de adquisición	Descripción y características
Imagen óptica multispectral	Satélite SPOT-5	29-08-2010	Banda 1: Verde (0,50 - 0,59 μm) Banda 2: Rojo (0,61 - 0,68 μm) Banda 3: Infrarrojo cercano (0,78 - 0,89 μm) Banda 4: Infrarrojo medio (1,58 - 1,75 μm)
Nubes de puntos LiDAR	Escáner láser aeroportado Leica ALS60	23-01-2011 y 02-01-2011	Densidad nominal: 0,5 puntos/m ² . Hasta 4 retornos por pulso sin clasificar. Precisión altimétrica superior a 0,20 m.
3 imágenes SAR	Satélites ENVISAT-1 / ERS-2	29-01-2011 / 17-02-2011 / 09-12-2010	Banda C: 5,6 cm

Tabla 1. Descripción de los datos de teledetección utilizados en el estudio.

III Congreso Nacional de I+D en Defensa y Seguridad, 2015

Entre julio y septiembre de 2014 se llevó a cabo una campaña de trabajo de campo con la finalidad de identificar *in situ* las siete tipologías de combustibles existentes según la clasificación de *Prometheus* (Tabla 2). Los datos fueron obtenidos en 108 parcelas de 15 m de radio (Figura 1), distribuidas mediante un muestreo aleatorio estratificado, tratando de lograr una muestra representativa de la variabilidad de la cobertura vegetal, su altura, la fracción de cabida cubierta y las pendientes topográficas de la zona de estudio. Las coordenadas del centro de cada parcela fueron tomadas con un GPS-GNSS modelo GS15 de Leica, alcanzando una precisión en planimetría de 0,33 m.

Tipo	Descripción del combustible y rangos de altura	Cobertura
T1	Pastizales agrícolas y vegetación herbácea.	Pasto >60%
T2	Pastizales y arbustos de 0,3-0,6 m con un 30-40% de herbáceas.	Matorral >60%; Arbolado <50%
T3	Arbustos de 0,6-2,0 m, así como árboles jóvenes resultantes de una regeneración natural o una reforestación.	
T4	Arbustos de 2,0-4,0 m y árboles regenerados.	
T5	Arbolado de más de 4 m sin sotobosque.	Arbolado >50%; Matorral <30%.
T6	Arbolado de más de 4 m con sotobosque donde la distancia de éste con la base del dosel arbóreo es superior a 0,5 m.	Arbolado >50%; Matorral >30%.
T7	Arbolado de más de 4 m con sotobosque alto y muy denso, donde la distancia de éste con la base del dosel arbóreo es inferior a 0,5 m.	

Tabla 2. Clasificación de tipos de combustible según *Prometheus*.

2.3. Procesamiento de las imágenes

En relación a las nubes de puntos LiDAR, fueron filtradas para separar los retornos del terreno del resto de información, utilizando el algoritmo de curvatura multiescala implementado en el software MCC 2.1 [10]. Los puntos clasificados como suelo fueron interpolados con el método “*Point-TIN-Raster*” para crear un modelo digital de elevaciones (MDE), con una resolución de 1 m, que permitiera normalizar la nube de puntos. Un conjunto de 29 métricas derivadas de los retornos láser, que presentan relación con la estructura forestal, fueron obtenidas utilizando el software FUSION LDV 3.30. Además, para lograr una mejor caracterización vertical de los combustibles, de acuerdo con Mutlu et al. [8], se generó una serie de bandas con información sobre el porcentaje de retornos LiDAR dentro de un rango específico de altura. Para la estimación de estos porcentajes se agruparon los puntos en las bandas que van de 0-0,5 m (BL1), 0,5-1 m (BL2), 1-2 m (BL3), 2-4 m (BL4), >4 m (BL5), 3-3,5 m (BL6), 3,5-4 m (BL7) y 1-4 m (BL8).

Por lo que respecta a la imagen óptica, fue procesada para obtener los valores de reflectividad y se calcularon los índices de vegetación NDVI (*Normalized Difference Vegetation Index*) y NDII (*Normalized Difference Infrared Index*) [11]. Las imágenes SAR fueron georreferenciadas utilizando un MDE remuestreado a 20 m a partir de los MDEs de 5 m disponibles en el Centro de Descargas del Centro Nacional de Información Geográfica (CNIG). Posteriormente, fueron calibradas radiométricamente de forma absoluta normalizando la retrodispersión con respecto a la unidad de superficie en la geometría terreno considerando el ángulo de incidencia local de cada píxel [11].

III Congreso Nacional de I+D en Defensa y Seguridad, 2015

Finalmente, se adoptó un tamaño de píxel común de 10 m, dado que corresponde con la resolución original de la imagen óptica y se consideró adecuado por la disponibilidad de suficientes retornos láser para calcular las métricas y bandas LiDAR. Además, los valores de todas las bandas susceptibles de ser utilizadas en el proceso de clasificación fueron reescaladas de 0 a 100.

2.4. Selección de la muestra de entrenamiento-validación y de las bandas originales

A partir de la localización de los centroides de las parcelas de campo, con apoyo en las ortofotografías del PNOA, se seleccionó un total de 2314 píxeles a los que se asignó el tipo de combustible. Asimismo, se seleccionaron 392 píxeles correspondientes a suelo desnudo. Posteriormente, el 10 % de esta muestra inicial, 267 píxeles, se tomaron aleatoriamente para la fase de validación y el resto se utilizó en la fase de entrenamiento.

Con el objeto de seleccionar las bandas más idóneas para la clasificación digital de tipos de combustible, se aplicó la prueba estadística no paramétrica de contraste de *Kruskal-Wallis*, que es adecuada cuando no es posible asumir la normalidad de la muestra.

2.5. Combinación de datos y clasificación digital

Se utilizaron dos métodos de combinación de información antes de llevar a cabo la posterior fase de clasificación [8]: i) análisis de componentes principales (ACP); y ii) fracción de ruido mínimo (MNF). Estos métodos permiten reducir la dimensionalidad o el número de bandas originales, derivándose nuevos factores (bandas) que explican la mayor cantidad de la variabilidad existente en las originales.

En cuanto a la clasificación digital multispectral, se utilizó el método paramétrico supervisado de máxima probabilidad empleando los píxeles de entrenamiento [11]. Finalmente, para verificar las imágenes clasificadas, se generaron las correspondientes matrices de confusión y porcentajes de acierto global (relación entre el número de píxeles correctamente asignados y el total), utilizando los píxeles de validación seleccionados previamente. Además se calculó el estadístico *kappa*, que mide la diferencia entre el acuerdo mapa-realidad observado y el que cabría esperar simplemente por azar [11].

3. Resultados y Discusión

Como se puede observar en la Tabla 3, las mejores bandas según el análisis *Kruskal-Wallis* fueron: i) el índice NDVI, que proporciona información sobre la cantidad, estado y desarrollo de la vegetación; ii) las bandas ópticas originales de la imagen SPOT-5; iii) todas las bandas LiDAR (BL) a excepción de aquellas cuyos rangos van de 0,5-1 m (BL2) y 1-2 m (BL3); y iv) las métricas derivadas de la nube de puntos relacionadas con la elevación media del dosel vegetal, el valor del percentil 75, la varianza o variabilidad de las alturas de la vegetación y el porcentaje de primeros retornos sobre la altura media, que expresaría el grado de cobertura vegetal. Por otro lado, las imágenes SAR presentaron valores muy bajos de *Chi-cuadrado* en comparación con el resto de bandas de información.

Tras realizar diversas combinaciones de bandas, tanto originales como transformadas, se presentan en la Tabla 4 los mejores resultados obtenidos, los cuales ponen de manifiesto, a través de la fiabilidad global de la clasificación supervisada, que la combinación de bandas ópticas con bandas LiDAR supera la precisión alcanzada utilizando únicamente las bandas de SPOT-5.

III Congreso Nacional de I+D en Defensa y Seguridad, 2015

Origen de la información	Bandas	Chi-cuadrado
Imagen óptica SPOT-5	NDVI	2000,8
Nube de puntos LiDAR	BL1	1993,1
	Elevación media de las alturas de los puntos (EM)	1993,0
	Percentil 75 (P ₇₅)	1974,9
	Varianza de las alturas de los puntos (V)	1949,2
Imagen óptica SPOT-5	Banda 4	1948,2
	Banda 2	1925,2
	Banda 1	1854,4
Nube de puntos LiDAR	BL5	1731,9
	BL8	1595,1
	BL4	1570,9
	BL7	1363,8
Imagen óptica SPOT-5	Banda 3	1318,6
Nube de puntos LiDAR	BL6	1317,7
	Porcentaje de primeros retornos sobre la altura media (%Ret)	1316,4
Imagen SAR ENVISAT-1	29-01-2011	441,5

Tabla 3. Selección de las mejores bandas de acuerdo con los valores de *Chi*-cuadrado resultantes de la prueba de *Kruskal-Wallis* con un nivel de significación estadística $P\text{-valor} \leq 0,01$.

Los resultados obtenidos están en la línea de los de otros autores [8,9], demostrando la idoneidad de la metodología aplicada para identificar tipos de combustible. Por otro lado, los métodos de reducción (ACP y MNF) no lograron superar el nivel de acierto alcanzado por el multibanda integrado por las bandas originales de SPOT-5, las bandas y métricas LiDAR (BL1,4,5,6,7,8; EM, P₇₅, V, %Ret) y el NDVI. La fiabilidad global de la clasificación fue del 72,7%, mientras que por tipo de combustible los porcentajes de acierto fueron: T1 75,0%, T2 70,0%, T3 50,0%, T4 88,9%, T5 71,4%, T6 55,6%, y T7 75,9%. El combustible T3 tiende a confundirse con el T2, al igual que ocurre entre el T6 y el T5. Esto puede ser debido a la baja densidad de puntos LiDAR, en comparación con estudios previos, que impide caracterizar con precisión las formaciones vegetales de menor porte (diferencias entre combustibles T2 y T3), e influye en la capacidad de penetración a través del dosel para detectar la presencia de sotobosque (diferencias entre combustibles T5 y T6).

Multibanda	Acierto global	k
Bandas SPOT-5	59,2 %	0,5
Bandas SPOT-5 + NDVI + BL1,4,5,6,7,8 + EM, P ₇₅ , V, %Ret	72,7 %	0,7
ACP componentes 1 a 9 (bandas SPOT-5 + NDVI + BL1,4,5,6,7,8 + EM, P ₇₅ , V, %Ret)	61,4 %	0,6
MNF componentes 1 a 8 (bandas SPOT-5 + BL1,4,5,6,7,8 + EM, P ₇₅ , V, %Ret) + NDVI	72,3 %	0,7

Tabla 4. Comparación de los mejores resultados derivados de la clasificación digital. Coeficiente *kappa* (k) con un nivel de significación estadística $P\text{-valor} \leq 0,01$.

En la Figura 1 se muestra el mapa resultante de la clasificación supervisada. Si se analiza con detalle, existen píxeles localizados en zonas de escarpe y barrancos encajados (sector suroeste del mapa) que, siendo en la realidad combustible T1 o T2, han sido clasificados como combustibles con más de un 50% de arbolado. Este efecto, susceptible de ser analizado con más detalle en el futuro,

III Congreso Nacional de I+D en Defensa y Seguridad, 2015

parece producirse como consecuencia de los cambios bruscos de gradiente topográfico, que afectan tanto a la precisión de la nube de puntos, como a la respuesta espectral captada por el sensor.

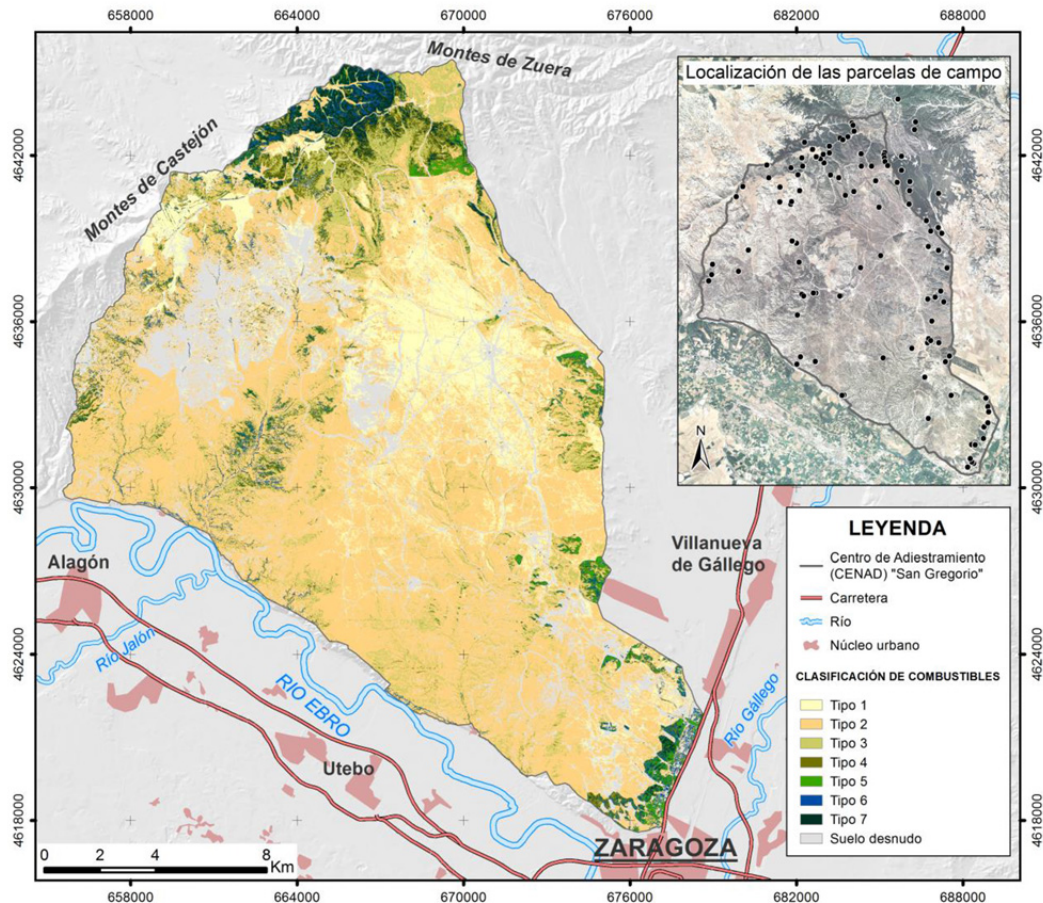


Figura 1. Mapa de combustibles del CENAD "San Gregorio".

4. Conclusiones

Este estudio demuestra cómo se puede mejorar la clasificación de combustibles combinando datos de teledetección pasiva y activa. La fiabilidad global de la cartografía presentada es aceptable, aunque es necesario seguir investigando sobre la incorporación de imágenes ópticas de mayor resolución o nuevas informaciones, por ejemplo relacionadas con la topografía de la superficie, con objeto de corregir los problemas derivados de una mala identificación del combustible en zonas de grandes cambios de pendiente. Asimismo, existen otros métodos de clasificación más novedosos, tales como los árboles de decisión (*random forest*) que pueden mejorar estos resultados. La obtención de una cartografía digital de alta resolución (10 m) de modelos de combustible permite su utilización en simuladores de incendio, facilitando en gran medida las posibles tareas de extinción y prevención en el CENAD.

III Congreso Nacional de I+D en Defensa y Seguridad, 2015

Agradecimientos

Este trabajo ha sido financiado por el Proyecto de Investigación del Centro Universitario de la Defensa de Zaragoza (Proyecto N° 2013-04) y la beca predoctoral (FPI BOA 30, 11/02/2011) del Gobierno de Aragón. Los autores agradecen al CNIG y al Centro de Información Territorial de Aragón por facilitar los datos LiDAR-PNOA y las imágenes SPOT-5, así como a la ESA por proporcionar los datos SAR (Proyecto N° 18850). Asimismo agradecen al CENAD “San Gregorio” su inestimable asistencia en los trabajos de campo.

Referencias

1. Montealegre, A.L.; Lamelas, M.T.; Tanase, M.; de la Riva, J. Forest Fire Severity Assessment Using ALS Data in a Mediterranean Environment. *Remote Sensing* **2014**, *6*, 4240-4265.
2. Ministerio de Defensa. Directiva 42/2010, de 12 de julio, de la Ministra de Defensa para la prevención y extinción de incendios forestales en los campos de maniobras y tiro. Disponible en: <http://www.defensa.gob.es/Galerias/areasTematicas/medioAmbiente/fichero/DGN-Directiva-42-sobre-Incendios-Forestales.pdf> (consultado 6 julio 2015).
3. BOE. Ley 21/2015, de 20 de julio, por la que se modifica la Ley 43/2003, de 21 de noviembre, de Montes. Boletín Oficial del Estado, 21 de julio de 2015, núm. 173, pp. 60234-60272.
4. Ramírez, J.; Blanco, J.; Cerrillo, A.; Marqués, G.; Valcarce, R.; Llordén, J.M.. Generación de cartografía de modelos de combustible a partir de datos LiDAR y análisis de imágenes orientado a objetos para su integración en Wildfire Analyst. En *Montes: Servicios y desarrollo rural. 6º Congreso Forestal Español*; Sociedad Española de ciencias forestales: Vitoria-Gasteiz, España, 2013; pp. 2–14.
5. Rothermel, R.C. A mathematical model for fire spread predictions in wildland fires. *USDA Forest Service Research Paper* **1972**, INT-115, 1-40.
6. PROMETHEUS S.V. Project. 2000. Management techniques for optimization of suppression and minimization of wildfire effects.
7. Riaño, D.; Chuvieco, E.; Salas, J.; Palacios-Orueta, A.; Bastarrika, A. Generation of fuel map type from Landsat TM images and ancillary data in Mediterranean ecosystems. *Canadian Journal of Forest Resources* **2002**, *32*, 1301-1315.
8. Mutlu, M.; Popescu, S.C.; Stripling, C.; Spencer, T. Mapping surface fuel models using lidar data and multispectral data fusion for fire behavior. *Remote Sensing of Environment* **2008**, *112*, 274-285.
9. Jakubowski, M.K.; Guo Q.; Collins, B.; Stephens, S.; Kelly, M. Predicting surface fuel models and fuel metrics using Lidar and CIR imagery in a dense, mountainous forest. *Photogrammetric Engineering & Remote Sensing* **2013**, *79*, 37–49.
10. Montealegre, A.L.; Lamelas, M.T.; de la Riva, J. A comparison of open source LiDAR filtering algorithms in a Mediterranean forest environment. *IEEE Trans. Sel. Top. Appl. Earth Observ. Remote Sens.* **2015**. En prensa doi:10.1109/JSTARS.2015.2436974.
11. Chuvieco, E. *Teledetección ambiental. La observación de la Tierra desde el espacio*; Ariel: Barcelona, España, 2010; p. 592.

4 CONCLUSIONES Y LÍNEAS DE TRABAJO FUTURAS

El último capítulo de la tesis está dedicado a la valoración general de las aportaciones realizadas sobre el uso de los datos LiDAR del PNOA en espacios forestales mediterráneos. Los resultados obtenidos permiten extraer una serie de conclusiones generales y específicas, que ya han sido discutidas en cada una de las publicaciones, y de las cuales se derivan unas líneas de trabajo futuras. Hay que recalcar que esta tesis doctoral ha abierto una nueva línea de investigación en el seno del Departamento de Geografía y Ordenación del Territorio de la Universidad de Zaragoza. Su carácter pionero favorece la utilización del LiDAR en el ámbito de la ciencia geográfica y abre nuevas puertas para el desarrollo de investigaciones cada vez más complejas.

4.1 Conclusiones principales

Durante los últimos veinte años, los sistemas LiDAR aeroportados se han utilizado en muchas disciplinas como una herramienta de recogida automática y eficiente de datos espaciales con coordenadas x , y , z . En la presente tesis se han presentado, desde un punto de vista crítico, distintas aplicaciones forestales, prestando especial atención al tratamiento previo de las nubes de puntos, como es el filtrado y la interpolación para generar MDE que permitan la normalización de las alturas de la vegetación o de cualquier otro elemento situado sobre la superficie.

Frente a otras aplicaciones que utilizan altas densidades de puntos y/o están desarrolladas en entornos forestales de bosque boreal o de estructura homogénea, el punto fuerte de esta investigación radica en el uso de datos de baja densidad de puntos en entornos forestales heterogéneos, donde el cuerpo de investigación está mucho menos desarrollado.

Los datos LiDAR del PNOA constituyen una nueva fuente de información topográfica de carácter público con un potencial todavía por explotar para múltiples aplicaciones. La metodología semi-empírica desarrollada en esta tesis ha puesto de manifiesto la necesidad de comparar distintos tipos de filtros para seleccionar cuál es el que mejor se adapta a las características de una nube de puntos en un área de ambiente mediterráneo. Como se ha podido comprobar, los diferentes filtros tienen sus ventajas e inconvenientes, ninguno de ellos es perfecto y todos son susceptibles de cometer errores en función de la complejidad de la escena. Por consiguiente, continúa siendo un reto seleccionar apropiadamente los parámetros de los filtros, ya que dependen del área de estudio y de los datos LiDAR empleados. No obstante, esta investigación permite orientar y guiar a los usuarios potenciales de nubes de puntos de densidad baja, como las del LiDAR-PNOA, para seleccionar el mejor filtro adaptado a un típico paisaje mediterráneo con una destacable diversidad de cubiertas y de características topográficas.

El filtro MCC, diseñado para ambientes forestales, es el que ha funcionado excepcionalmente bien, ya que maximiza la cantidad de puntos de suelo e incrementa el detalle de la superficie del terreno.

El matorral rebrotador, los tocones y las ramas son las cubiertas más problemáticas para el filtrado, así como las pendientes superiores a 15° . Sin embargo, conclusiones menos firmes pueden derivarse del efecto de la densidad de puntos y del ángulo de escaneo, porque los filtros morfológicos rompen la tendencia al ser menos sensibles a estos factores.

Una de las principales debilidades de los filtros morfológicos es la dificultad para mantener ciertas características de la topografía cuando las ventanas de filtrado cambian iterativamente en las operaciones de apertura y cierre, siendo crítica la selección del tamaño de dichas ventanas. A ello se suma la complejidad en su parametrización.

Llama la atención cómo el filtro de FUSION, que es un *software* especialmente diseñado para aplicaciones forestales, no es el que ha proporcionado mejores resultados, sobre todo por la cierta dificultad que implica la selección de los parámetros del algoritmo.

La investigación ha demostrado que la mayor parte de los filtros evaluados pueden identificar puntos de suelo relativamente bien, aunque la densidad de puntos sea baja. Dicho de otra manera, en densidades superiores a 2 puntos/m², el filtro también encuentra dificultades, por lo que el error en el filtrado está más influenciado por la complejidad de la escena que por la densidad de puntos.

En definitiva, es necesario desarrollar un completo procedimiento de filtrado aplicable a la nube de puntos LiDAR-PNOA, aunque por el momento los métodos basados en superficies parametrizadas

ráster o TIN parecen ser la mejor alternativa para el tipo de paisaje analizado.

Por otro lado, no existe un método de interpolación que sea universalmente mejor que los demás. Cada interpolador produce predicciones usando diferentes suposiciones acerca de los datos, por lo que un MDE puede dar cuenta de la variación local mejor que otro en función de diversos factores. En general, la precisión de los MDE ha sido mayor con la resolución espacial de 1 m, lo que indica que la elección del tamaño del píxel puede ser más importante que la elección del tipo de interpolador. Además, las diferencias entre resoluciones son más importantes en terrenos que presentan mayor complejidad topográfica y cubierta vegetal.

Del análisis de la influencia de diferentes variables en los errores de interpolación se desprende que todas causan un aumento en los errores medios absolutos en determinadas condiciones, como por ejemplo valores elevados de pendiente y bajas densidades de puntos. El método TIN a ráster resultó ser el menos sensible a la influencia de estas variables explicativas analizadas, si bien es cierto que todos los interpoladores presentaron precisiones similares tanto en términos globales, como en las estratificaciones por factores influyentes en el error.

Se ha evidenciado que la densidad de retornos de suelo, que es dependiente de la configuración del sistema LiDAR y de la cobertura vegetal existente en la zona, la resolución espacial del ráster y la complejidad morfológica del terreno afectan a la capacidad de los interpoladores para generar MDE precisos. En este sentido, se puede concluir que el método TIN a ráster es el más preciso en términos generales y además cuenta con la ventaja de que no tiene parámetros ajustables complejos, como sucede con el *kriging*, que puedan introducir sesgos por parte del usuario.

El análisis CART indica que la variabilidad topográfica y la densidad de puntos tienen una influencia significativa en el MDE, pero también la cubierta, que determina la cantidad de retornos que alcanzan el suelo. A medida que aumenta la complejidad del terreno también lo hace la incertidumbre en el error del MDE.

La investigación desarrollada también puede orientar, en la selección de métodos de interpolación, a los usuarios de nubes de puntos de baja densidad en ausencia de información que describa la idoneidad de parámetros de interpolación en áreas ocupadas por bosque de pino junto con matorrales perennifolios.

Se ha podido demostrar la utilidad de los datos de baja densidad de puntos para el inventario de masas forestales de pino carrasco. Los coeficientes de determinación y los errores obtenidos están dentro de los habitualmente alcanzados en otros estudios, siendo peor en todos ellos la estimación de la densidad de pies. Es relevante señalar que, si bien algunos estudios han logrado mejores ajustes y/o el cumplimiento de los supuestos estadísticos de los modelos de regresión lineal transformando logarítmicamente las variables originales, en este trabajo tan solo ha sido necesario para la estimación del ratio de copa. En general, los percentiles incluidos en los modelos se relacionan de forma directa con la altura media, el diámetro y el volumen, al igual que ocurre con los porcentajes de primeros retornos sobre 1 m, en el caso del área basimétrica y el ratio de copa, o el porcentaje de todos los retornos sobre 1 m en los modelos de volumen, densidad de pies y biomasa. Por el contrario, la asimetría de la distribución presenta una relación inversa con el área basimétrica, ya que cuanto menor es aquélla, más concentrados están los retornos en torno a valores altos. En el caso del coeficiente de variación, cuanto más elevado es, mayor heterogeneidad existe entre las alturas de los retornos, lo cual se relaciona negativamente con el tamaño de los árboles y con la densidad de pies. Por otro lado, si la distribución de las alturas de los retornos es leptocúrtica, es indicador de una masa desarrollada, de ahí que se relacione positivamente con el volumen y con la biomasa.

Los resultados obtenidos sobre el contenido de carbono de la biomasa aérea ponen de manifiesto la

utilidad de los datos LiDAR, ya que a partir de las ecuaciones alométricas de biomasa –tras aplicar los pertinentes coeficientes de conversión a carbono– puede cartografiarse dicha variable.

Con respecto a la estimación de la severidad post-incendio combinando variables LiDAR, que informan de la estructura de la vegetación tras el paso del fuego, y la valoración que se extrae del índice CBI, la regresión logística constituye una alternativa adecuada –cuando la muestra no sigue una distribución normal– para conocer la probabilidad de que exista un nivel de severidad alto o bajo con una precisión relativamente aceptable. Se trata de una aplicación novedosa que surge ante las limitaciones de los sensores ópticos para conocer los daños estructurales de la vegetación. Para desarrollar métodos operacionales robustos y aplicables en entornos heterogéneos respecto al tipo de vegetación y a las condiciones ambientales es necesario entender la relación existente entre la información procedente de los retornos láser y las características de las superficies quemadas. Por ello, la investigación se ha centrado en correlacionar y contrastar ambas informaciones con objeto de conocer qué variables LiDAR son las más adecuadas para el análisis de la severidad. Hay que tener en cuenta, no obstante, las limitaciones que implica utilizar un índice sintético como el CBI, que tiene un cierto componente subjetivo. Además, no se ha contado con información LiDAR previa al incendio para poder calcular cuantitativamente el cambio estructural pre- y post-fuego experimentado por la vegetación.

Por último, teniendo en cuenta que el Ministerio de Defensa ha venido demostrando a lo largo de los últimos años un fuerte compromiso por la protección medioambiental, trabajando a través de planes y programas en la conservación de los terrenos que gestiona, la investigación desarrollada ha propuesto una metodología para mejorar la identificación y caracterización de combustibles forestales combinando datos de teledetección pasiva y activa. La obtención de una cartografía digital de alta resolución (10 m) basada en el modelo *Prometheus* permite caracterizar los tipos de combustibles existentes en el CENAD “San Gregorio”, para servir de ayuda a las tareas de prevención y extinción. Cabe destacar que la clasificación digital obtenida se puede integrar en un SIG para producir, mediante métodos de simulación, predicciones que ayuden a los expertos a determinar despliegues de medios y refuerzos en los lugares con mayor riesgo.

La fiabilidad global de la clasificación obtenida es aceptable, aunque es necesario seguir investigando sobre la incorporación de imágenes ópticas de mayor resolución o incluso imágenes SAR. El uso de estas últimas en medios mediterráneos y en temáticas específicamente relacionadas con incendios está muy restringido debido a la inherente heterogeneidad orográfica del territorio, no habiéndose encontrado apenas referencias de su uso para la obtención de cartografía de combustibles, con lo que éste se erige como un campo abierto a la investigación.

En definitiva, esta tesis trata, en parte, de dar respuesta al objetivo planteado por el proyecto LiDAR-PNOA, relacionado con el fomento de la utilización de estos datos en ámbitos multidisciplinares, buscando mejoras en el procesado de los datos, para reducir costes y tiempos en la generación de productos útiles.

4.2 Conclusiones específicas

A continuación se recogen las conclusiones que se derivan de los objetivos específicos que se detallan en el Capítulo 1:

Identificar el filtro o método de clasificación que mejor se adapta a las características de la nube de puntos LiDAR-PNOA en un ambiente forestal heterogéneo a través de una metodología semi-empírica, mediante la comparación de los implementados en software libre.

- A nivel cuantitativo el filtro MCC presentó el error global más bajo (16,7%), mientras que MLS y ETW cometieron los mayores errores (37,5% y 33,7%, respectivamente). El error tipo II en MCC fue del 20,8%, en comparación con el 12,7% de error tipo I. La inclinación a cometer errores tipo II no debería ser un gran inconveniente de cara a utilizar esa nube de puntos filtrada, teniendo en cuenta que este tipo de errores pueden ser editados manualmente mucho más fácilmente que los errores tipo I. De los siete filtros comparados, MCC es el que mejor equilibra estos dos tipos de errores.
- El segundo mejor filtro en relación con el porcentaje de acierto global es LAStools, con un 82,8%. La diferencia con MCC radica en que solo un 13,7% de los puntos objeto fueron erróneamente clasificados como suelo, mientras que un 20,8% de los puntos fueron rechazados como suelo.
- Los filtros MCC y LAStools son los más adecuados para los usuarios que quieran obtener resultados precisos con una mínima parametrización del algoritmo. No obstante, conviene señalar que el primero es computacionalmente más intenso y el tiempo de ejecución podría resultar un inconveniente a la hora de procesar nubes de puntos de grandes dimensiones.
- MLS es el algoritmo que ha alcanzado el mayor porcentaje de error tipo I (75,0%) frente a un 0% de error tipo II, lo cual indica que el filtro ha descartado una gran cantidad de puntos de suelo. No obstante, PM fue el mejor de la categoría de filtros morfológicos.
- Las mayores dificultades para los filtros analizados se ocasionan cuando la pendiente topográfica supera los 15°, debido a las distorsiones que produce cuando los algoritmos aplican umbrales de altura para diferenciar los puntos de suelo del resto. De hecho, los filtros implementados en ALDPAT fueron los que más errores de clasificación produjeron en las zonas de orografía compleja. Por el contrario, MCC y LAStools parecen ser los filtros que mejor toleran la presencia de puntos en pendientes mayores de 15°.
- En cuanto al tipo de cubierta forestal, el filtro MCC funcionó significativamente mejor que el resto de filtros comparados, especialmente ante matorrales y zonas ocupadas por arbustos rebrotadores y troncos, tocones y ramas dispuestas sobre el terreno (restos de un incendio). En la cubierta de bosque de coníferas, MLS, BCAL y LAStools presentan los mejores resultados. Precisamente este último, es el que se revela como más adecuado para filtrar escenas dominadas por pastizales y campos de cultivo, al no encontrarse errores en esta categoría.
- Conclusiones poco firmes pueden extraerse del efecto de la densidad de puntos en el filtrado, ya que los filtros analizados no presentan un patrón claro que relacione un mayor error con una menor cantidad de puntos disponibles para el filtrado.
- En cuanto al ángulo de escaneo, tampoco se pueden extraer grandes conclusiones, ya que aunque el error tiende a disminuir cuando el ángulo de escaneo es superior a $\pm 14^\circ$, existen algunas excepciones: PM y ETW parecen ser menos sensibles a los errores introducidos por la presencia de puntos con ángulos de escaneo superiores a ese umbral.

Analizar la idoneidad de diversos métodos de interpolación para crear MDE de alta resolución con los retornos de suelo desnudo, dado que son necesarios para la normalización de las alturas de la nube de puntos y, por ende, de la vegetación en el contexto de las aplicaciones forestales del LiDAR.

- El análisis de la calidad interna de los MDE revela que el método de interpolación TIN a ráster es el más preciso para generar MDE de 1 m de resolución espacial. Por el contrario, el método punto a ráster ha reportado la peor precisión, pese a estar implementado en un *software* específicamente diseñado para aplicaciones forestales.

- El análisis de la calidad externa de los MDE indica mayores RMSE entre (37 y 63 cm), siendo el interpolador IDW el más adecuado para generar MDE de 1 m de paso de malla.
- Con el propósito de analizar el efecto de la pendiente, la cobertura del terreno, la densidad de puntos de suelo y el ratio de penetración de pulsos, se concluye que el método de interpolación TIN a ráster es la solución más adecuada cuando la escena presenta una amplia variabilidad de pendientes topográficas, la densidad es inferior a 0,5 puntos/m² y las cubiertas son complejas (matorral, bosque de pino, arbustos densos y de bajo porte, tocones y ramas sobre el suelo, etc.)
- El análisis CART muestra que las áreas con una combinación de pendientes elevadas por encima de 15° y baja densidad de puntos (inferior a 0,3 puntos/m²) son las más propensas a presentar mayores errores en la interpolación.

*Estimar variables dasométricas en masas de *P. halepensis* Mill., así como el contenido de carbono de estas masas empleando modelos de regresión lineal múltiple que relacionan las variables derivadas de los retornos láser y los datos de campo.*

- Los modelos de regresión lineal ajustados para estimar la altura media, el diámetro cuadrático medio, el área basimétrica, el volumen, la densidad de la masa, el ratio de copa y la biomasa presentan un R² de 0,87, 0,84, 0,89, 0,89, 0,48, 0,69 y 0,89, respectivamente con RMSE de 0,72 m, 1,99 cm, 2,39 m²/ha, 11,01 m³/ha, 187,53 pies/ha, 7,26% y 11,418.11 kg/ha respectivamente. Todos ellos evidencian, a excepción de la densidad de pies de pino, que se pueden estimar con precisión variables dasométricas en masas de *P. halepensis*.
- En la mayor parte de los modelos se han incluido variables LiDAR de tres tipos: relacionadas con la altura del dosel arbóreo (percentiles y elevación máxima), variables que muestran la variabilidad de la altura del dosel vegetal (desviación estándar, asimetría, curtosis y coeficiente de variación de la distribución de los retornos) y variables relacionadas con la densidad del dosel (proporción de retornos por encima de un umbral de altura sobre el terreno).

Analizar las relaciones que existen entre los indicadores de la severidad del fuego estimados en campo y las variables derivadas de las nubes de puntos mediante un ajuste de regresión logística para obtener una cartografía de cuatro áreas quemadas.

- Se presenta una aproximación metodológica novedosa basada en un análisis de regresión logística para relacionar la información estructural que proporcionan las variables derivadas de la distribución de las alturas de los retornos láser con los datos de severidad estimados en campo mediante el índice CBI.
- Las variables independientes LiDAR curtosis o el percentil 25 de las alturas de los retornos presentan las correlaciones más altas con los valores de CBI, mientras que la distinción entre la clase de severidad alta (CBI>1,5) y baja (CBI<1,5) se puede apreciar mejor con el percentil 20 y con el porcentaje de todos los retornos por encima de la altura media.
- El modelo de regresión logística está constituido por dos variables LiDAR: *canopy relief ratio* y el porcentaje total de retornos por encima de 1 m de altura sobre el terreno, que informan sobre la distribución y la densidad de ramas y hojas del dosel.
- La cartografía generada sobre la severidad del fuego en los incendios de Zuera, Jaulín, Aliaga y Los Olmos –derivada del modelo de regresión logística– proporciona un apoyo inicial para valorar el impacto causado por el fuego sobre los pinares.
- El modelo de regresión logística obtenido tiende a subestimar las zonas de moderada-alta severidad, pero presenta la mejor precisión en las zonas de severidad más alta, donde precisamente es crítica la intervención por parte de los gestores con objeto de mitigar las

consecuencias del incendio sobre el ecosistema.

- Las correlaciones entre el valor de CBI de las parcelas de campo y los habitualmente utilizados índices espectrales, dNBR y RdNBR, resultaron levemente superiores (R^2 de 0,7 y 0,68, respectivamente) en comparación con la correlación de los valores de CBI y los estimados por el modelo de regresión logística (R^2 de 0,63). No obstante, hay que tener en cuenta que el modelo ajustado no contempla información sobre la situación pre-incendio.

Desarrollar una metodología para obtener cartografía de tipos de combustible forestal mediante la combinación de información LiDAR-PNOA e imágenes multiespectrales de alta resolución.

- La combinación de bandas espectrales, un índice de vegetación como el NDVI y bandas LiDAR relacionadas con la distribución de las alturas de los retornos láser, supera la precisión alcanzada por la clasificación digital utilizando únicamente las bandas de la imagen SPOT 5.
- Las técnicas para reducir la dimensionalidad del multibanda (ACP y MNF) han revelado resultados sensiblemente peores en la clasificación digital de los combustibles. La diversidad de bandas (variables) consideradas en el análisis puede suponer una dificultad a la hora de extraer la información relevante.
- La fiabilidad global de la cartografía de combustibles presentada es aceptable (73%) y está en la línea de los resultados obtenidos por otros autores que han seguido una aproximación basada en la combinación de bandas de sensores ópticos y activos. Sin embargo, a nivel de tipos de combustibles se observan algunas imprecisiones. Las mayores confusiones se producen entre los tipos 3 y 2 (arbustos hasta 2 m y pastizales) y entre 6 y 5 (arbolado sin y con sotobosque). También hay ciertos errores en las zonas de topografía accidentada, especialmente en barrancos de incisión lineal y en zonas de escarpe.

4.3 Líneas de trabajo futuras

La investigación presentada ha profundizado en la aplicación de los datos públicos LiDAR-PNOA, que son de baja densidad, en espacios forestales mediterráneos, para proporcionar una primera “guía” sobre el método de filtrado e interpolación más adecuados y mostrar algunas de las posibilidades que tienen para la estimación de variables dasométricas, analizar la severidad y cartografiar los tipos de combustible forestal con apoyo en trabajo de campo. Puede decirse que esta investigación constituye la “semilla” o el inicio de otras investigaciones todavía más prometedoras, existiendo algunas posibles líneas de mejora a considerar en futuros trabajos.

A continuación se resumen diferentes propuestas de trabajo, algunas de las cuales están en curso:

- En relación con el filtrado, sería interesante analizar con más detalle los errores tipo II (inclusión de puntos pertenecientes a objetos de la escena en la categoría de suelo desnudo), puesto que son los que suelen producir abultamientos u otro tipo de artefactos irreales en los MDE. Esto estaría en relación con las cubiertas ocupadas por matorral denso y de escasa altura donde el filtro encuentra dificultades para clasificar los puntos.
- Sería conveniente testar los algoritmos de filtrado analizados, además de otros más recientes, en otro tipo de ambientes forestales; por ejemplo, en bosques de caducifolias (siempre y cuando las fechas de vuelo correspondan con la época de mayor vigor vegetal), para aumentar el conocimiento sobre su uso y, principalmente, determinar los parámetros más idóneos para llevar a cabo una clasificación lo más precisa posible. En esta línea, teniendo presente que la parametrización del algoritmo depende de la escena y de los datos de entrada, es interesante

seguir con este tipo de investigaciones que sirvan de “guía” a los usuarios.

- Es de interés analizar la propagación del error en el filtrado sobre el MDE, así como la influencia de éste sobre los modelos de regresión obtenidos mediante un modelo empírico que, además, espacializase el error o nivel de incertidumbre.
- Otros métodos de clasificación digital, como los no paramétricos, podrían utilizarse para comprobar si superan en precisión al método supervisado de máxima verosimilitud. En este sentido, a tenor de los resultados obtenidos, convendría considerar otras fuentes de información, como por ejemplo un MDE de la zona de estudio u otras variables relacionadas con la orografía, con el objetivo de mejorar, si se puede, la identificación de los combustibles situados en zonas escarpadas.
- Se podrían ajustar otro tipo de modelos, por ejemplo *random forest*, para estimar las mismas variables dasométricas y comparar los resultados. La literatura científica consultada evidencia buenos resultados tanto con los métodos de regresión lineal como con los no paramétricos.
- Naturalmente, convendría ahondar en la aplicación de los datos LiDAR-PNOA para generar modelos para otro tipo de especies forestales del ámbito mediterráneo, no solo coníferas. Además, es crucial generar modelos para superficies de matorral, que es una vía de trabajo poco explorada.
- Convendría profundizar en la utilidad de los datos LiDAR para estimar biomasa residual, incluso combinando estos datos con información espectral o SAR.
- Se podría ahondar en el conocimiento de la masa boscosa estudiada generando un índice de complejidad estructural utilizando las variables dasométricas estimadas.
- Sería deseable comparar los modelos de estimación de las variables dasométricas utilizando la segunda cobertura del PNOA, la cual estará disponible en un futuro muy próximo.
- Se podrían utilizar imágenes ópticas de las áreas incendiadas para realizar una aproximación basada en la combinación de bandas de información para desarrollar un modelo robusto y aplicable para estimar la severidad. Incluso, si se dispusiera de nubes de puntos pre- y post-fuego, podrían generarse índices a partir de las métricas de los retornos láser.
- Ya que el interés por generar modelos de estimación radica en su generalización para superficies amplias, se podrían explorar técnicas de reducción de la densidad de puntos para facilitar un tratamiento y manejo de las nubes mucho más eficiente. En ocasiones, dependiendo de las características del terreno y del método utilizado, es posible reducir la cantidad de datos LiDAR para lograr un proceso eficiente.

REFERENCIAS

- Abdullahi, S., Schardt, M., & Pretzsch, H. (2017). An unsupervised two-stage clustering approach for forest structure classification based on X-band InSAR data — A case study in complex temperate forest stands. *International Journal of Applied Earth Observation and Geoinformation*, 57, 36–48. doi:10.1016/j.jag.2016.12.010
- Ackermann, F. (1999). Airborne laser scanning: present status and future expectations. *ISPRS Journal of Photogrammetry and Remote Sensing*, 54, 64–67. doi:10.1016/s0924-2716(99)00009-x
- Aguilar, F.J., Agüera, F., Aguilar, M.A., & Carvajal, F. (2005). Effects of terrain morphology, sampling density, and interpolation methods on grid DEM accuracy. *Photogrammetric Engineering and Remote Sensing*, 71, 805–816. doi: 10.14358/PERS.71.7.805
- Aguilar, F.J., & Mills, J.P. (2008). Accuracy assessment of LiDAR-derived digital elevation models. *The Photogrammetric Record*, 23, 148–169. doi:10.1111/j.1477-9730.2008.00476.x
- Alberti, G., Boscutti, F., Pirotti, F., Bertacco, C., De Simon, G., Sigura, M., Cazorzi, F., & Bonfanti, P. (2013). A LiDAR-based approach for a multi-purpose characterization of Alpine forests: an Italian case study. *IForest - Biogeosciences and Forestry*, 6, 156-168. doi:10.3832/ifor0876-006
- Albini, F.A. (1976). *Estimating wildfire behavior and effects*. Ogden, Utah: Dept. of Agriculture, Forest Service, Intermountain Forest and Range Experiment Station, General Technical Report INT-30, 92 pp.
- Amato, V.J.W., Lightfoot, D., Stropki, C., & Pease, M. (2013). Relationships between tree stand density and burn severity as measured by the Composite Burn Index following a ponderosa pine forest wildfire in the American Southwest. *Forest Ecology and Management*, 302, 71–84. doi:10.1016/j.foreco.2013.03.015
- Andersen, H., McGaughey, R., & Reutebuch, S.E. (2005). Estimating forest canopy fuel parameters using LiDAR data. *Remote Sensing of Environment*, 94, 441–449. doi:10.1016/j.rse.2004.10.013
- Andersen, H., Reutebuch, S., & McGaughey, R. (2006). A rigorous assessment of tree height measurements obtained using airborne LiDAR and conventional field methods. *Canadian Journal of Remote Sensing* 32, 355–366. doi:10.5589/m06-030
- Anderson, E.S., Thompson, J.A., Crouse, D.A., & Austin, R.E. (2006). Horizontal resolution and data density effects on remotely sensed LiDAR-based DEM. *Geoderma*, 132, 406–415. doi:10.1016/j.geoderma.2005.06.004
- Arroyo, L.A., Pascual, C., & Manzanera, J.A. (2008). Fire models and methods to map fuel types: the role of remote sensing. *Forest Ecology and Management*, 256, 1239–1252.

doi:10.1016/j.foreco.2008.06.048

- ASPRS (2016). *LASeR File Format Exchange Activities* [en línea]. [Consulta: 12 de diciembre de 2016]. Disponible en web: <<https://www.asprs.org/committee-general/laser-las-file-format-exchange-activities.html>>
- Axelsson, P. (1999). Processing of laser scanner data-algorithms and applications. *ISPRS Journal of Photogrammetry and Remote Sensing*, 54, 138-147. doi:10.1016/S0924-2716(99)00008-8
- Axelsson, P. (2000). DEM generation from laser scanner data using adaptive TIN models. *International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, 33, 110-117.
- Baltsavias, E.P. (1999a). A comparison between photogrammetry and laser scanning. *ISPRS Journal of Photogrammetry and Remote Sensing*, 54, 83-94. doi:10.1016/S0924-2716(99)00014-3
- Baltsavias, E.P. (1999b). Airborne laser scanning: basic relations and formulas. *ISPRS Journal of Photogrammetry and Remote Sensing*, 54, 199-214. doi:10.1016/S0924-2716(99)00015-5
- Baltsavias, E.P. (1999c). Airborne laser scanning: existing systems and firms and other resources. *ISPRS Journal of Photogrammetry and Remote Sensing*, 54, 164-198. doi:10.1016/S0924-2716(99)00016-7
- Baskerville, G.L. (1972). Use of logarithmic regression in the estimation of plant biomass. *Canadian Journal of Forest Research*, 2, 49-53. doi:10.1139/x72-009
- Bater, C.W., & Coops, N.C. (2009). Evaluating error associated with LiDAR-derived DEM interpolation. *Computers and Geosciences*, 35, 289-300. doi:10.1016/j.cageo.2008.09.001
- Beguet, B., Boukir, S., Guyon, D., & Chehata, N. (2013). Modelling-based feature selection for classification of forest structure using very high resolution multispectral imagery. In: *IEEE International Conference on Systems, Man, and Cybernetics*. Manchester, pp. 4294-4299. doi:10.1109/SMC.2013.732
- Behan, A. (2000). On the matching accuracy of rasterized scanning laser altimeter data. *International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, 33, 75-82.
- Berlanga, V., Rubio, M.J., & Vilà, R. (2013). Cómo aplicar árboles de decisión en SPSS. *Revista d'Innovació i Recerca en Educació*, 6, 65-79. doi:10.1344/reire2013.6.1615
- Bolton, D.K., Coops, N.C., & Wulder, M.A. (2015). Characterizing residual structure and forest recovery following high-severity fire in the western boreal of Canada using Landsat time-series and airborne LiDAR data. *Remote Sensing of Environment*, 163, 48-60. doi:10.1016/j.rse.2015.03.004
- Botalico, F., Chirici, G., Giannini, R., Mele, S., Mura, M., Puxeddu, M., McRoberts, R.E., Valbuena, R., & Travaglini, D. (2017). Modeling Mediterranean forest structure using airborne laser scanning data. *International Journal of Applied Earth Observation and Geoinformation* 57, 145-153. doi:10.1016/j.jag.2016.12.013
- Bouvier, M., Durrieu, S., Fournier, R.A., & Renaud, J.-P. (2015). Generalizing predictive models of forest inventory attributes using an area-based approach with airborne LiDAR data. *Remote Sensing of Environment*, 156, 322-334. doi:10.1016/j.rse.2014.10.004
- Breiman, L., Friedman, J., Stone, C.J., & Olshen, R.A. (1984). *Classification and Regression Trees*, The Wadsworth and Brooks-Cole statistics-probability series. Taylor & Francis, Pacific Grove, California, 368 pp.
- Burgan, R.E., Klaver, R.W., & Klarer, J.M. (1998). Fuel models and fire potential from satellite and surface observations. *International Journal of Wildland Fire*, 8, 159-170.

- Campbell, J.B. (2006). *Introduction to remote sensing*. Taylor & Francis, Oxon, Reino Unido, 626 pp.
- Carter, J., Schmid, K., Waters, K., Betzhold, L., Hadley, B., Mataosky, R., & Halleran, J. (2012). *Lidar 101: An Introduction Lidar Technology, Data, and Applications*. National Oceanic and Atmospheric Administration (NOAA) Coastal Services Center, Charleston, 72 pp.
- Caruso, C., & Quarta, F. (1998). Interpolation methods comparison. *Computers and Mathematics with Applications*, 35, 109–126. doi:10.1186/cc2857.
- Castillo-Santiago, M.A., Ricker, M., & de Jong, B.H.J. (2010). Estimation of tropical forest structure from SPOT-5 satellite images. *International Journal of Remote Sensing*, 31, 2767–2782. doi:10.1080/01431160903095460
- Céline, E., Philippe, M., Astrid, V., Catherine, B., Musampa, C., & Pierre, D. (2013). National forest cover change in Congo Basin: Deforestation, reforestation, degradation and regeneration for the years 1990, 2000 and 2005. *Global Change Biology*, 19, 1173–1187. doi:10.1111/gcb.12092
- CNIG (2017). Centro Nacional de Información Geográfica. <https://www.cnig.es/> [Consulta: viernes, 24 de marzo de 2017].
- Chasmer, L., Hopkinson, C., & Treitz, P. (2006). Investigating laser pulse penetration through a conifer canopy by integrating airborne and terrestrial LiDAR. *Canadian Journal of Remote Sensing*, 32, 116–125. doi:10.5589/m06-011
- Chen, Q., Gong, P., Baldocchi, D., & Xie, G. (2007). Filtering airborne laser scanning data with morphological methods. *Photogrammetric Engineering and Remote Sensing* 73, 175–185. doi:10.14358/pers.73.2.175
- Chuvieco, E., Martín, M.P. (2004). *Nuevas tecnologías para la estimación del riesgo de incendios forestales*. Colección de Estudios Ambientales y Socioeconómicos. Consejo Superior de Investigaciones Científicas, Instituto de Economía y Geografía, Madrid, 190 pp.
- Chuvieco, E. (2009). *Earth observation of wildland fires in Mediterranean ecosystems*. Springer, Alcalá de Henares, 257 pp.
- Chuvieco, E. (2010). *Teledetección ambiental. La observación de la Tierra desde el espacio*. Ariel Ciencia, Barcelona, 594 pp..
- Coops, N., Hilker, T., Wulder, M., St-Onge, B., Newnham, G., Siggins, A., & Trofymow, J. (2007). Estimating canopy structure of Douglas-fir forest stands from discrete-return LiDAR. *Trees - Structure and Function*, 21, 295–310–310. doi:10.1007/s00468-006-0119-6
- Cressie, N. (1988). Spatial prediction and ordinary kriging. *Mathematical Geology*, 20, 405–421. doi:10.1007/BF00892986
- Cuadrat, J.M. (2004). El clima de Aragón. In: J.L., Peña, L.A., Longares, & M. Sánchez (Eds.), *Geografía Física de Aragón: Aspectos generales y temáticos*. Universidad de Zaragoza e Institución Fernando el Católico, Zaragoza, pp. 15–26.
- De Santis, A., & Chuvieco, E. (2009). GeoCBI: A modified version of the Composite Burn Index for the initial assessment of the short-term burn severity from remotely sensed data. *Remote Sensing of Environment*, 113, 554–562. doi:10.1016/j.rse.2008.10.011
- Edwards, A.C., Maier, S.W., Hutley, L.B., Williams, R.J., & Russell-Smith, J. (2013). Spectral analysis of fire severity in north Australian tropical savannas. *Remote Sensing of Environment*, 136, 56–65. doi:10.1016/j.rse.2013.04.013
- Ene, L.T., Næsset, E., Gobakken, T., Gregoire, T.G., Ståhl, G., & Nelson, R. (2012). Assessing the

- accuracy of regional LiDAR-based biomass estimation using a simulation approach. *Remote Sensing of Environment*, 123, 579–592. doi:10.1016/j.rse.2012.04.017
- Erdody, T.L., & Moskal, L.M. (2010). Fusion of LiDAR and imagery for estimating forest canopy fuels. *Remote Sensing of Environment*, 114, 725–737. doi:10.1016/j.rse.2009.11.002
- Estornell, J., Ruiz, L.A., Velázquez-Martí, B., & Hermosilla, T. (2012). Estimation of biomass and volume of shrub vegetation using LiDAR and spectral data in a Mediterranean environment. *Biomass Bioenergy*, 46, 710–721. doi:10.1016/j.biombioe.2012.06.023
- Evans, J., & Hudak, A. (2007). A multiscale curvature algorithm for classifying discrete return LiDAR in forested environments. *IEEE Transactions on Geoscience and Remote Sensing*, 45, 1029–1038. doi:10.1109/TGRS.2006.890412
- Evans, J., Hudak, A., Faux, R., & Smith, A.M. (2009). Discrete return LiDAR in natural resources: Recommendations for project planning, data processing, and deliverables. *Remote Sensing*, 1, 776–794. doi:10.3390/rs1040776
- García, M., Riaño, D., Chuvieco, E., & Danson, F.M. (2010). Estimating biomass carbon stocks for a Mediterranean forest in central Spain using LiDAR height and intensity data. *Remote Sensing of Environment*, 114, 816–830. doi:10.1016/j.rse.2009.11.021
- García, M., Riaño, D., Chuvieco, E., Salas, J., & Danson, F.M. (2011). Multispectral and LiDAR data fusion for fuel type mapping using Support Vector Machine and decision rules. *Remote Sensing of Environment*, 115, 1369–1379. doi:10.1016/j.rse.2011.01.017
- García, M.J.L., & Caselles, V. (1991). Mapping burns and natural reforestation using thematic Mapper data. *Geocarto International*, 6, 31–37. doi:10.1080/10106049109354290
- Gatzliolis, D., & Andersen, H. (2008). *A guide to LIDAR data acquisition and processing for the forests of the Pacific Northwest*. Department of Agriculture, Forest Service, Pacific Northwest Research Station. Portland, 32 pp.
- Geoportal IDEE (2017). <http://www.idee.es/europeo-inspire> [Consulta: viernes, 24 de marzo de 2017].
- Glenn, E.P., Huete, A.R., Nagler, P.L., & Nelson, S.G. (2008). Relationship between remotely-sensed vegetation indices, canopy attributes and plant physiological processes: What vegetation indices can and cannot tell us about the landscape. *Sensors*, 8, 2136–2160. doi:10.3390/s8042136
- Gómez, C., Wulder, M.A., Montes, F., & Delgado, J.A. (2012). Modeling forest structural parameters in the Mediterranean pines of central Spain using QuickBird-2 imagery and Classification and Regression Tree Analysis (CART). *Remote Sensing*, 4, 135–159. doi:10.3390/rs4010135
- Gonçalves-Seco, L., González-Ferreiro, E., Diéguez-Aranda, U., Fraga-Bugallo, B., Crecente, R., & Miranda, D. (2011). Assessing the attributes of high-density *Eucalyptus globulus* stands using airborne laser scanner data. *International Journal of Remote Sensing*, 32, 9821–9841. doi:10.1080/01431161.2011.593583
- Gong, J., Li, Z., Zhu, Q., Sui, H., & Zhou, Y. (2000). Effects of various factors on the accuracy of DEMs: An intensive experimental investigation. *Photogrammetric Engineering and Remote Sensing*, 66, 1113–1117.
- González-De Vega, S., De las Heras, J., & Moya, D. (2016). Resilience of Mediterranean terrestrial ecosystems and fire severity in semiarid areas: Responses of Aleppo pine forests in the short, mid and long term. *Science of The Total Environment*, 573, 1171–1177. doi:10.1016/j.scitotenv.2016.03.115

- Gonzalez-Ferreiro, E., Dieguez-Aranda, U., & Miranda, D. (2012). Estimation of stand variables in *Pinus radiata* D. Don plantations using different LiDAR pulse densities. *Forestry*, 85, 281–292. doi:10.1093/forestry/cps002
- González-Olabarria, J.-R., Rodríguez, F., Fernández-Landa, A., & Mola-Yudego, B. (2012). Mapping fire risk in the Model Forest of Urbión (Spain) based on airborne LiDAR measurements. *Forest Ecology and Management*, 282, 149–156. doi:10.1016/j.foreco.2012.06.056
- Green, A.A., Berman, M., Switzer, P., & Craig, M.D. (1988). A transformation for ordering multispectral data in terms of image quality with implications for noise removal. *IEEE Transactions on Geoscience and Remote Sensing*, 26, 65–74. doi:10.1109/36.3001
- Guerra-Hernández, J., Tomé, M., & González-Ferreiro, E. (2016). Using low density LiDAR data to map Mediterranean forest characteristics by means of an area-based approach and height threshold analysis. *Revista de Teledetección*, 46, 103–117. doi:10.4995/raet.2016.3980.
- Guo, Q., Li, W., Yu, H., & Alvarez, O. (2010). Effects of topographic variability and LiDAR sampling density on several DEM interpolation methods. *Photogrammetric Engineering and Remote Sensing*, 76, 701–712. doi: 10.14358/PERS.76.6.701
- Hair, J.F., Anderson, R.E., Tatham, R.L., & Black, W.C. (1999). *Análisis multivariante*. Prentice Hall Iberia, Madrid, 832 pp.
- Hall, S.A., Burke, I.C., Box, D.O., Kaufmann, M.R., & Stoker, J.M. (2005). Estimating stand structure using discrete-return LiDAR: an example from low density, fire prone ponderosa pine forests. *Forest Ecology and Management*, 208, 189–209. doi:10.1016/j.foreco.2004.12.001
- Hanley, J.A., & McNeil, B. (1982). The meaning and use of the área under a Receiver Operating Characteristic (ROC) curve. *Radiology*, 143, 29–36. doi:10.1148/radiology.143.1.7063747
- Hill, J.M., Graham, L.A., & Henry, R.J. (2000). Wide-area topographic mapping and applications using airborne light detection and ranging (LiDAR) technology. *Photogrammetric Engineering and Remote Sensing*, 66, 908–914.
- Hodgson, M.E., & Shelley, B.M. (1994). Removing the topographic effect in remotely sensed imagery. In: J., Kirn (Ed.), *Erdas Monitor*. Erdas Incorporated, Atlanta, pp. 4-6.
- Hodgson, M.E., Jensen, J., Raber, G., Tullis, J., Davis, B.A., Thompson, G., & Schuckman, K. (2005). An evaluation of LiDAR-derived elevation and terrain slope in leaf-off conditions. *Photogrammetric Engineering and Remote Sensing*, 71, 817–823. doi:10.14358/PERS.71.7.817
- Hopkinson, C.C., L.E., Zsigovics, G., Creed, I.F., Sitar, M., Treitz, P., Maher, R.V., 2004. Errors in LiDAR ground elevation and wetland vegetation height estimates. *International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences*, 36, 108–113.
- Hu, H., Ding, Y., Zhu, Q., Wu, B., Lin, H., Du, Z., Zhang, Y., & Zhang, Y. (2014). An adaptive surface filter for airborne laser scanning point clouds by means of regularization and bending energy. *ISPRS Journal of Photogrammetry and Remote Sensing*, 92, 98–111. doi:10.1016/j.isprsjprs.2014.02.014
- Huising, E.J., & Gomes Pereira, L.M. (1998). Errors and accuracy estimates of laser data acquired by various laser scanning systems for topographic applications. *ISPRS Journal of Photogrammetry and Remote Sensing*, 53, 245–261. doi:10.1016/s0924-2716(98)00013-6
- Hunt, E.R., & Rock, B.N. (1989). Detection of changes in leaf water content using Near- and Middle-Infrared reflectances. *Remote Sensing of Environment*, 30, 43–54. doi:10.1016/0034-4257(89)90046-1

- Hutchinson, M.F. (1989). A new procedure for gridding elevation and stream line data with automatic removal of spurious pits. *Journal of Hydrology*, 106, 211–232. doi:10.1016/0022-1694(89)90073-5
- Hutchinson, M.F., Xu, T., & Stein, A. (2011). Recent progress in the ANUDEM elevation gridding procedure. In: T. Hengel, I.S. Evans, J.P. Wilson & M. Gould (Eds.), *Geomorphometry*, Redlands, California, pp. 19–22.
- Hyde, P., Dubayah, R., Walker, W., Blair, J.B., Hofton, M., & Hunsaker, C. (2006). Mapping forest structure for wildlife habitat analysis using multi-sensor (LiDAR, SAR/InSAR, ETM+, Quickbird) synergy. *Remote Sensing of Environment*, 102, 63–73. doi:10.1016/j.rse.2006.01.021
- Hyypä, J., Hyypä, H., Litkey, P., Yu, X., Haggrén, H., Rönnholm, P., Pyysalo, U., Pitkänen, J., & Maltamo, M. (2004). Algorithms and methods of airborne laser scanning for forest measurements. *International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences*, 36, 82–89.
- Hyypä, H., Yu, X., Hyypä, J., Kaartinen, H., Kaasalainen, S., Honkavaara, E., & Rönnholm, P. (2005). Factors affecting the quality of DTM generation in forested areas. *International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, 36, 85–90.
- Hyypä, J., Yu, X., Hyypä, H., & Maltamo, M. (2006). Methods of airborne laser scanning for forest information extraction. In: T., Koukal, & W., Schneider (Eds.), *Proceedings of the Workshop on 3D Remote Sensing in Forestry*. Viena, pp. 1–16.
- IBM SPSS (2011). *IBM SPSS Decision Trees 20* [en línea]. [Consulta: 12 de diciembre de 2016]. Disponible en web: http://portal.uned.es/pls/portal/docs/PAGE/UNED_MAIN/BIBLIOTECA/APOYO%20ESTUDIANTES/GUIAGENERALBIBLIOTECA_9.PDF
- Jakubowski, M.K., Guo, Q., Collins, B., Stephens, S., Kelly, M. (2013). Predicting surface fuel models and fuel metrics using LiDAR and CIR imagery in a dense mixed conifer forest. *Photogrammetric Engineering and Remote Sensing*, 79, 37–49. doi:10.14358/PERS.79.1.37
- Jordán de Urriés, F., & Ureta, M.I. (2009). La restauración de los montes de Zuera y Castejón de Valdejasa. *Natural de Aragón*, 36, 10–11.
- Kane, V.R., Lutz, J.A., Roberts, S.L., Smith, D.F., McGaughey, R.J., Povak, N.A., & Brooks, M.L. (2013). Landscape-scale effects of fire severity on mixed-conifer and red fir forest structure in Yosemite National Park. *Forest Ecology and Management*, 287, 17–31. doi:10.1016/j.foreco.2012.08.044
- Kane, V.R., North, M.P., Lutz, J.A., Churchill, D.J., Roberts, S.L., Smith, D.F., McGaughey, R.J., Kane, J.T., & Brooks, M.L. (2014). Assessing fire effects on forest spatial structure using a fusion of landsat and airborne LiDAR data in Yosemite national park. *Remote Sensing of Environment*, 151, 89–101. doi:10.1016/j.rse.2013.07.041
- Kankare, V., Vastaranta, M., Holopainen, M., Rätty, M., Yu, X., Hyypä, J., Hyypä, H., Alho, P., & Viitala, R. (2013). Retrieval of forest aboveground biomass and stem volume with airborne scanning LiDAR. *Remote Sensing*, 5, 2257–2274. doi:10.3390/rs5052257
- Kasischke, E.S., Turetsky, M.R., Ottmar, R.D., French, N.H.F., Hoy, E.E., & Kane, E.S. (2008). Evaluation of the composite burn index for assessing fire severity in Alaskan black spruce forests. *International Journal of Wildland Fire*, 17, 515–526. doi:10.1071/WF08002
- Keeley, J.E., Brennan, T., & Pfaff, A.H. (2008). Fire severity and ecosystem responses following crown fires in California shrublands. *Ecological Applications*, 18, 1530–1546. doi:10.1890/07-0836.1
- Key, C.H., & Benson, N. (2003). *The Composite Burn Index (CBI): Field rating of burn severity*. Department

- of the Interior, US Geological Survey, Northern Rocky Mountain Science Center.
- Key, C.H., & Benson, N.C. (2006). Landscape assessment (LA). In: D.C., Lutes, R.E., Keane, J.F., Caratti, C.H., Key, N.C., Benson, S., Sutherland & L.J., Gangi (Eds.), *FIREMON: Fire effects monitoring and inventory system*. Department of Agriculture, Forest Service, Rocky Mountain Research Station, Fort Collins, pp. 1-55.
- Kienzle, S. (2004). The effect of DEM raster resolution on first order, second order and compound terrain derivatives. *Transactions in GIS*, 8, 83–111. doi:10.1111/j.1467-9671.2004.00169.x
- Kilian, J., Haala, N., & Englich, M. (1996). Capture and evaluation of Airborne Laser Scanner data. *International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, 31, 383–388.
- Kobler, A., Pfeifer, N., Ogrinc, P., Todorovski, L., Ostir, K., & Dzeroski, S. (2007). Repetitive interpolation: A robust algorithm for DTM generation from Aerial Laser Scanner Data in forested terrain. *Remote Sensing of Environment*, 108, 9–23. doi:10.1016/j.rse.2006.10.013
- Koetz, B., Morsdorf, F., van der Linden, S., Curt, T., & Allgöwer, B. (2008). Multi-source land cover classification for forest fire management based on imaging spectrometry and LiDAR data. *Forest Ecology and Management*, 256, 263–271. doi:10.1016/j.foreco.2008.04.025
- Kramer, H.A., Collins, B.M., Kelly, M., & Stephens, S.L. (2014). Quantifying ladder fuels: A new approach using LiDAR. *Forests*, 5, 1432–1453. doi:10.3390/f5061432
- Kraus, K., & Pfeifer, N. (1998). Determination of terrain models in wooded areas with airborne laser scanner data. *ISPRS Journal of Photogrammetry and Remote Sensing*, 53, 193–203. doi:10.1016/s0924-2716(98)00009-4
- Larose, D.T. (2005). *Discovering knowledge in data: An introduction to data mining*. John Wiley & Sons, Hoboken, 222 pp.
- Lasaponara, R., & Lanorte, A. (2007). Remotely sensed characterization of forest fuel types by using satellite ASTER data. *International Journal of Applied Earth Observation and Geoinformation*, 9, 225–234. doi:10.1016/j.jag.2006.08.001
- Lasaponara, R., Lanorte, A., Pignatti, S. (2006). Characterization and mapping of fuel types for the Mediterranean ecosystems of Pollino National Park in Southern Italy by using hyperspectral MIVIS Data. *Earth Interactions*, 10, 1–11. doi:10.1175/EI165.1
- Latifi, H., Nothdurft, A., & Koch, B. (2010). Non-parametric prediction and mapping of standing timber volume and biomass in a temperate forest: application of multiple optical/LiDAR-derived predictors. *Forestry*, 83, 395–407. doi:10.1093/forestry/cpq022
- Latifi, H., Fassnacht, F.E., Müller, J., Tharani, A., Dech, S., & Heurich, M. (2015). Forest inventories by LiDAR data: A comparison of single tree segmentation and metric-based methods for inventories of a heterogeneous temperate forest. *International Journal of Applied Earth Observation and Geoinformation*, 42, 162–174. doi:10.1016/j.jag.2015.06.008
- Lefsky, M.A., Cohen, W.B., Parker, G.G., & Harding, D.J. (2002). LiDAR remote sensing for ecosystem studies. *BioScience*, 52, 19-30. doi:10.1641/0006-3568(2002)052[0019:LRSFES]2.0.CO;2
- Lentile, L.B., Holden, Z.A., Smith, A.M.S., Falkowski, M.J., Hudak, A.T., Morgan, P., Lewis, S.A., Gessler, P.E., & Benson, N.C. (2006). Remote sensing techniques to assess active fire characteristics and post-fire effects. *International Journal of Wildland Fire*, 15, 319-345. doi:10.1071/WF05097

- Lim, K., Treitz, P., Wulder, M., St-Onge, B., & Flood, M. (2003). LiDAR remote sensing of forest structure. *Progress in Physical Geography*, 27, 88-106. doi: 10.1191/0309133303pp360ra
- Li, Z., Zhu, Q., & Gold, C. (2004). *Digital terrain modeling: Principles and methodology*. CRC Press, USA, 340 pp.
- Li, Y., Wu, H., Xu, H., An, R., Xu, J., & He, Q., 2013. A gradient-constrained morphological filtering algorithm for airborne LiDAR. *Optics and Laser Technology*, 54, 288–296. doi:10.1016/j.optlastec.2013.06.007
- Lindberg, E., & Hollaus, M. (2012). Comparison of methods for estimation of stem volume, stem number and basal area from Airborne Laser Scanning Data in a Hemi-Boreal forest. *Remote Sensing*, 4, 1004–1023. doi:10.3390/rs4041004
- Liu, W., Song, C., Schroeder, T.A., & Cohen, W.B. (2008). Predicting forest successional stages using multitemporal Landsat imagery with forest inventory and analysis data. *International Journal of Remote Sensing*, 29, 3855–3872. doi:10.1080/01431160701840166
- Liu, X. (2011). Accuracy assessment of LiDAR elevation data using survey marks. *Survey Review*, 43, 80–93. doi:10.1179/003962611X12894696204704
- Lloyd, C.D., & Atkinson, P.M. (2002). Deriving DSMs from LiDAR data with kriging. *International Journal of Remote Sensing*, 23, 2519–2524. doi:10.1080/01431160110097998
- Lloyd, C.D., & Atkinson, P.M. (2006). Deriving ground surface digital elevation models from LiDAR data with geostatistics. *International Journal of Geographical Information Science*, 20, 535–563. doi:10.1080/13658810600607337
- Longares, J.L. (2004). Variedad biogeográfica del territorio aragonés. In: J.L., Peña, L.A., Longares, & M. Sánchez (Eds.), *Geografía Física de Aragón: Aspectos generales y temáticos*. Universidad de Zaragoza e Institución Fernando el Católico, Zaragoza, pp. 27–40.
- López del Ríos, R., & Bardají, M. (2009). Estrategia de extinción en los grandes incendios del verano. *Natural de Aragón*, 35, 32–33.
- Maguya, A.S., Junttila, V., & Kauranne, T. (2013). Adaptive algorithm for large scale dtm interpolation from LiDAR data for forestry applications in steep forested terrain. *ISPRS Journal of Photogrammetry and Remote Sensing*, 85, 74–83. doi:10.1016/j.isprsjprs.2013.08.005
- Maguya, A.S., Tegel, K., Junttila, V., Kauranne, T., Korhonen, M., Burns, J., Leppanen, V., & Sanz, B. (2015). Moving voxel method for estimating canopy base height from airborne laser scanner data. *Remote Sensing*, 7, 8950–8972. doi:10.3390/rs70708950
- Maltamo, M., Næsset, E., & Vauhkonen, J. (2014). *Forestry Applications of Airborne Laser Scanning: Concepts and Case Studies*. Springer, London, 464 pp.
- Márquez, J. (2004). *Modelos digitales de elevaciones, métodos de interpolación y parámetros morfométricos del relieve*. Directores: J. Ojeda. Tesis doctoral. Universidad de Sevilla, Sevilla.
- McGaughey, R. (2009). *FUSION/LDV: Software for LiDAR data analysis and visualization*. US Department of Agriculture, Forest Service, Pacific Northwest Research Station, Seattle, 115 pp.
- Means, J., Acker, S., Fitt, B., Renslow, M., Emerson, L., & Hendrix, C. (2000). Predicting forest stand characteristics with airborne scanning LiDAR. *Photogrammetric Engineering and Remote Sensing*, 66, 1367–1371.
- Menard, S. (2010). *Logistic regression: From introductory to advanced concepts and applications*. SAGE, USA, 392 pp.

- Meng, X., Currit, N., & Zhao, K. (2010). Ground filtering algorithms for Airborne LiDAR Data: A review of critical issues. *Remote Sensing*, 2, 833–860. doi:10.3390/rs2030833
- Miller, J.D., Knapp, E.E., Key, C.H., Skinner, C.N., Isbell, C.J., Creasy, R.M., & Sherlock, J.W. (2009). Calibration and validation of the relative differenced Normalized Burn Ratio (RdNBR) to three measures of fire severity in the Sierra Nevada and Klamath Mountains, California, USA. *Remote Sensing of Environment*, 113, 645–656. doi:10.1016/j.rse.2008.11.009
- Miller, J.D., & Thode, A.E. (2007). Quantifying burn severity in a heterogeneous landscape with a relative version of the delta Normalized Burn Ratio (dNBR). *Remote Sensing of Environment*, 109, 66–80. doi:10.1016/j.rse.2006.12.006
- Ministerio de Defensa (2007). Centros de Adiestramiento y Campos de Maniobras y Tiro. *Ejército*, 800, 3–116.
- Mongus, D., & Žalik, B. (2012). Parameter-free ground filtering of LiDAR data for automatic DTM generation. *ISPRS Journal of Photogrammetry and Remote Sensing*, 67, 1–12. doi:10.1016/j.isprsjprs.2011.10.002
- Montorio, R., Pérez-Cabello, F., García-Martín, A., Vlassova, L., & de la Riva, J. (2014). La severidad del fuego: revisión de conceptos, métodos y efectos ambientales. In: *Geoecología, cambio ambiental y paisaje: homenaje al profesor José María García Ruiz*. Instituto Pirenaico de Ecología, Universidad de La Rioja, pp. 427–440.
- Morsdorf, F., Kötz, B., Meier, E., Itten, K.I., & Allgöwer, B. (2006). Estimation of LAI and fractional cover from small footprint airborne laser scanning data based on gap fraction. *Remote Sensing of Environment*, 104, 50–61. doi:10.1016/j.rse.2006.04.019
- Mutlu, M., Popescu, S.C., Stripling, C., & Spencer, T. (2008). Mapping surface fuel models using LiDAR and multispectral data fusion for fire behavior. *Remote Sensing of Environment*, 112, 274–285. doi:10.1016/j.rse.2007.05.005
- Naesset, E. (1997). Estimating timber volume of forest stands using airborne laser scanner data. *Remote Sensing of Environment*, 61, 246–253. doi: 10.1016/S0034-4257(97)00041-2
- Naesset, E. (2002). Predicting forest stand characteristics with airborne scanning laser using a practical two-stage procedure and field data. *Remote Sensing of Environment*, 80, 88–99. doi:10.1016/S0034-4257(01)00290-5
- Næsset, E., & Gobakken, T. (2008). Estimation of above- and below-ground biomass across regions of the boreal forest zone using airborne laser. *Remote Sensing of Environment*, 112, 3079–3090. doi:10.1016/j.rse.2008.03.004
- Naesset, E., & Okland, T. (2002). Estimating tree height and tree crown properties using airborne scanning laser in a boreal nature reserve. *Remote Sensing of Environment*, 79, 105–115. doi:10.1016/S0034-4257(01)00243-7
- Nord-Larsen, T., & Schumacher, J. (2012). Estimation of forest resources from a country wide laser scanning survey and national forest inventory data. *Remote Sensing of Environment*, 119, 148–157. doi:10.1016/j.rse.2011.12.022
- Pardo, A., & Ruíz, M.A. (2005). *Análisis de datos con SPSS 13 Base*. Mc Graw Hill, Madrid, 600 pp.
- Parks, S.A., Dillon, G.K., & Miller, C. (2014). A new metric for quantifying burn severity: The Relativized Burn Ratio. *Remote Sensing*, 6, 1827–1844. doi:10.3390/rs6031827
- Patenaude, G., Hill, R.A., Milne, R., Gaveau, D.L.A., Briggs, B.B.J., & Dawson, T.P. (2004). Quantifying

- forest above ground carbon content using LiDAR remote sensing. *Remote Sensing of Environment*, 93, 368–380. doi:10.1016/j.rse.2004.07.016
- Pausas, J.G., Llovet, J., Rodrigo, A., & Vallejo, R. (2008). Are wildfires a disaster in the Mediterranean basin? – A review. *International Journal of Wildland Fire*, 17, 713–723. doi:10.1071/WF07151
- Peña, J.L., & Lozano, M.V. (2004). Las unidades del relieve aragonés. In: J.L., Peña, L.A., Longares, & M. Sánchez (Eds.), *Geografía Física de Aragón: Aspectos generales y temáticos*. Universidad de Zaragoza e Institución Fernando el Católico, Zaragoza, pp. 3–14.
- Pfeifer, N., Stadler, P., & Briese, C. (2001). Derivation of digital terrain models in the SCOP++ environment. In: *Proceedings of the OEEPE Workshop on Airborne Laserscanning and Interferometric SAR for Digital Elevation Models*. Stockholm, pp. 1–13.
- Pfeifer, M., Gonsamo, A., Disney, M., Pellikka, P., & Marchant, R. (2012). Leaf area index for biomes of the Eastern Arc Mountains: Landsat and SPOT observations along precipitation and altitude gradients. *Remote Sensing of Environment*, 118, 103–115. doi:10.1016/j.rse.2011.11.009
- Pingel, T.J., Clarke, K.C., & McBride, W.A. (2013). An improved simple morphological filter for the terrain classification of airborne LIDAR data. *ISPRS Journal of Photogrammetry and Remote Sensing*, 77, 21–30. doi:10.1016/j.isprsjprs.2012.12.002
- Platt, R.V. (2014). Wildfire hazard in the home ignition zone: An object-oriented analysis integrating LiDAR and VHR satellite imagery. *Applied Geography*, 51, 108–117. doi:10.1016/j.apgeog.2014.03.011
- Polat, N., & Uysal, M. (2015). Investigating performance of Airborne LiDAR data filtering algorithms for DTM generation. *Measurement*, 63, 61–68. doi:10.1016/j.measurement.2014.12.017
- Polat, N., Uysal, M., & Toprak, A.S. (2015). An investigation of DEM generation process based on LiDAR data filtering, decimation, and interpolation methods for an urban area. *Measurement*, 75, 50–56. doi:10.1016/j.measurement.2015.08.008
- Popescu, S.C., & Zhao, K. (2008). A voxel-based lidar method for estimating crown base height for deciduous and pine trees. *Remote Sensing of Environment*, 112, 767–781. doi:10.1016/j.rse.2007.06.011
- Prometheus (1999). *Management techniques for optimization of suppression and minimization of wildfire effects. System validation*. European Commission, DG XII, ENVIR & CLIMATE, Contract number ENV4-CT98-0716, 1998–2000.
- Ramírez, J., Blanco, J., Cerrillo, A., Marqués, G., Valcarce, R., & Llordén, J.M. (2013). Generación de cartografía de modelos de combustible a partir de datos LiDAR y análisis de imágenes orientado a objetos para su integración en Wildfire Analyst. In: Sociedad Española de Ciencias Forestales (Ed.), *Montes: Servicios y desarrollo rural. Actas del 6º Congreso Forestal Español*, Vitoria-Gasteiz, pp. 2–14.
- Renslow, M. (2013). *Manual of airborne topographic LiDAR*. The American Society for Photogrammetry and Remote Sensing, Bethesda, 504 pp.
- Reutebuch, S.E., McGaughey, R.J., Andersen, H.-E., & Carson, W.W. (2003). Accuracy of a high-resolution lidar terrain model under a conifer forest canopy. *Canadian Journal of Remote Sensing*, 29, 527–535. doi:10.5589/m03-022
- Reutebuch, S.E., Andersen, H.-E., & McGaughey, R.J. (2005). Light Detection and Ranging (LiDAR): an emerging tool for multiple resource inventory. *Journal of Forestry*, 103, 286–292.

- Riaño, D., Chuvieco, E., Salas, J., Palacios-Orueta, A., Bastarrika, A. (2002). Generation of fuel type maps from Landsat TM images and ancillary data in Mediterranean ecosystems. *Canadian Journal of Forest Research*, 32, 1301–1315. doi:10.1139/x02-052
- Riaño, D., Meier, E., Allgöwer, B., Chuvieco, E., & Ustin, S.L. (2003). Modeling airborne laser scanning data for the spatial generation of critical forest parameters in fire behavior modeling. *Remote Sensing of Environment*, 86, 177–186. doi:10.1016/s0034-4257(03)00098-1
- Riaño, D., Chuvieco, E., Condés, S., González-Matesanz, J., & Ustin, S.L. (2004a). Generation of crown bulk density for *Pinus sylvestris* L. from lidar. *Remote Sensing of Environment*, 92, 345–352. doi:10.1016/j.rse.2003.12.014
- Riaño, D., Valladares, F., Condés, S., & Chuvieco, E. (2004b). Estimation of leaf area index and covered ground from airborne laser scanner (LiDAR) in two contrasting forests. *Agricultural and Forest Meteorology*, 124, 269–275. doi:10.1016/j.agrformet.2004.02.005
- Riaño, D., Chuvieco, E., Ustin, S.L., Salas, J., Rodríguez-Pérez, J.R., Ribeiro, L.M., Viegas, D.X., Moreno, J.M., & Fernández, H. (2007). Estimation of shrub height for fuel-type mapping combining airborne LiDAR and simultaneous color infrared ortho imaging. *International Journal of Wildland Fire*, 16, 341–348. doi:10.1071/WF06003
- Richardson, J.J., Moskal, L.M., & Kim, S.-H. (2009). Modeling approaches to estimate effective leaf area index from aerial discrete-return LiDAR. *Agricultural and Forest Meteorology*, 149, 1152–1160. doi:10.1016/j.agrformet.2009.02.007
- Roggero, M. (2001). Airborne laser scanning: Clustering in raw data. *International Archives of Photogrammetry and Remote Sensing*, 34, 227–232.
- Rosette, J., Suárez, J., Nelson, R., Los, S., Cook, B., & North, P. (2012). LiDAR remote sensing for biomass assessment. In: T., Fatoyinbo (Ed.), *Remote sensing of biomass: Principles and applications*. InTech, Rijeka, Croatia, pp. 3–26.
- Rothermel, R.C. (1972). A mathematical model for predicting fire spread in wildland fuels. *USDA Forest Service research paper*. INT-115, 1–40.
- Rouse, J.W., Haas, R.H., Schell, J.A., Deering, D.W. & Harlan, J.C. (1974). Monitoring the vernal advancement of retrogradation of natural vegetation. Type III, Final Report, NASA/GSFC, Greenbelt, 371 pp.
- Ruiz-Peinado, R., del Rio, M., & Montero, G. (2011). New models for estimating the carbon sink capacity of Spanish softwood species. *Forest Systems*, 20, 176–188. doi:10.5424/fs/2011201-11643
- Sambridge, M., Braun, J., & McQueen, H. (1995). Geophysical parametrization and interpolation of irregular data using natural neighbours. *Geophysical Journal International*, 122, 837–857. doi:10.1111/j.1365-246X.1995.tb06841.x
- Shan, J., & Toth, C.K. (2008). *Topographic Laser Ranging and Scanning: Principles and Processing*. CRC Press, Boca Raton, 616 pp.
- Sibson, R. (1981). A brief description of Natural Neighbour interpolation. In: V. Barnett (Ed.), *Interpreting Multivariate Data*. John Wiley & Sons, New York, pp. 21–36.
- Sierra, J. (2013). Protección civil y Fuerzas Armadas. In: Ministerio de Defensa, Instituto Español de Estudios Estratégicos (Eds.), *España ante las emergencias y catástrofes. Las Fuerzas Armadas en colaboración con las autoridades civiles*. Cuadernos de Estrategia. Ministerio de Defensa, Madrid, pp. 69–98.

- Silván-Cárdenas, J.L., & Wang, L. (2006). A multi-resolution approach for filtering LiDAR altimetry data. *ISPRS Journal of Photogrammetry and Remote Sensing*, 61, 11–22. doi:10.1016/j.isprsjprs.2006.06.002
- Sithole, G. (2001). Filtering of laser altimetry data using a slope adaptive filter. *International Archives of Photogrammetry and Remote Sensing*, 34, 203–210.
- Sithole, G., & Vosselman, G. (2004). Experimental comparison of filter algorithms for bare-Earth extraction from airborne laser scanning point clouds. *ISPRS Journal of Photogrammetry and Remote Sensing*, 59, 85–101. doi:10.1016/j.isprsjprs.2004.05.004
- Smith, S.L., Holland, D.A., & Longley, P.A. (2004). The importance of understanding error in LiDAR digital elevation models. *International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, 35, 996–1001.
- Smreček, R., & Danihelová, Z. (2013). Forest stand height determination from low point density airborne laser scanning data in Roznava Forest enterprise zone (Slovakia). *IForest - Biogeosciences and Forestry*, 6, 48–54. doi:10.3832/ifor0767-006
- Solomon, S., Qin, D., Manning, M., Chen, Z., Marquis, M., Averyt, K., Tignor, M., & Miller, H.L. (2007). Climate change 2007: The physical science basis. Contribution of Working Group I to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change. Cambridge, United Kingdom and New York.
- Soverel, N.O., Perrakis, D.D.B., & Coops, N.C. (2010). Estimating burn severity from Landsat dNBR and RdNBR indices across western Canada. *Remote Sensing of Environment*, 114, 1896–1909. doi:10.1016/j.rse.2010.03.013
- Sprugel, D.G. (1983). Correcting for bias in log-transformed allometric equations. *Ecology*, 64, 209–210. doi:10.2307/1937343
- Stephens, P.R., Kimberley, M.O., Beets, P.N., Paul, T.S.H., Searles, N., Bell, A., Brack, C., & Broadley, J. (2012). Airborne scanning LiDAR in a double sampling forest carbon inventory. *Remote Sensing of Environment*, 117, 348–357. doi:10.1016/j.rse.2011.10.009
- Stereńczak, K., & Kozak, J. (2011). Evaluation of digital terrain models generated in forest conditions from airborne laser scanning data acquired in two seasons. *Scandinavian Journal of Forest Research*, 26, 374–384. doi:10.1080/02827581.2011.570781
- Sterenczak, K., Ciesielski, M., Balazy, R., & Zawila-Niedzwiecki, T. (2016). Comparison of various algorithms for DTM interpolation from LiDAR data in dense mountain forests. *European Journal of Remote Sensing*, 49, 599–621. doi:10.5721/EuJRS20164932
- Streutker, D.R., & Glenn, N.F. (2006). LiDAR measurement of sagebrush steppe vegetation heights. *Remote Sensing of Environment*, 102, 135–145. doi:10.1016/j.rse.2006.02.011
- Su, J., & Bork, E. (2006). Influence of vegetation, slope and LiDAR sampling angle on DEM accuracy. *Photogrammetric Engineering and Remote Sensing*, 72, 1265–1274. doi: 10.14358/PERS.72.11.1265
- Sulaiman, N.S., Majid, Z., & Setan, H. (2010). DTM generation from LiDAR data by using different filters in open-source software. *Geoinformation Science Journal*, 10, 89–109.
- Tanase, M.A., Santoro, M., Wegmüller, U., de la Riva, J., & Pérez-Cabello, F. (2010). Properties of X-, C- and L-band repeat-pass interferometric SAR coherence in Mediterranean pine forests affected by fires. *Remote Sensing of Environment*, 114, 2182–2194. doi:10.1016/j.rse.2010.04.021
- Tanase, M.A., de la Riva, J., & Pérez-Cabello, F. (2011). Estimating burn severity at the regional level

- using optically based indices. *Canadian Journal of Forest Research*, 41, 863–872. doi:10.1139/x11-011
- Tanase, M.A., Panciera, R., Lowell, K., Tian, S., García-Martín, A., & Walker, J.P. (2014). Sensitivity of L-Band Radar backscatter to forest biomass in semiarid environments: A comparative analysis of parametric and nonparametric models. *IEEE Transactions on Geoscience and Remote Sensing*, 52, 4671–4685. doi:10.1109/TGRS.2013.2283521
- Tanase, M.A., Kennedy, R., & Aponte, C. (2015). Radar Burn Ratio for fire severity estimation at canopy level: An example for temperate forests. *Remote Sensing of Environment*, 170, 14–31. doi:10.1016/j.rse.2015.08.025
- Thompson, J.A., Bell, J.C., & Butler, C.A. (2001). Digital elevation model resolution: Effects on terrain attribute calculation and quantitative soil-landscape modeling. *Geoderma*, 100, 67–89. doi:10.1016/S0016-7061(00)00081-1
- Tian, X., Douglas J. M., Shu, L., & Wang, M. (2005). Fuel classification and mapping from satellite images. *Journal of Forest Research*, 16, 311–316. doi:10.1007/bf02858198
- Varghese, A.O., Suryavanshi, A., & Joshi, A.K. (2016). Analysis of different polarimetric target decomposition methods in forest density classification using C band SAR data. *International Journal of Remote Sensing*, 37, 694–709. doi:10.1080/01431161.2015.1136448
- Vega, J.A., Fernández, C., Pérez-Gorostiaga, P., & Fonturbel, T. (2008). The influence of fire severity, serotiny, and post-fire management on *Pinus pinaster* Ait. recruitment in three burnt areas in Galicia (NW Spain). *Forest Ecology and Management*, 256, 1596–1603. doi:10.1016/j.foreco.2008.07.005
- Veraverbeke, S., Hook, S., & Hulley, G. (2012). An alternative spectral index for rapid fire severity assessments. *Remote Sensing of Environment*, 123, 72–80. doi:10.1016/j.rse.2012.02.025
- Verbyla, D.L., Kasischke, E.S., & Hoy, E.E. (2008). Seasonal and topographic effects on estimating fire severity from Landsat TM/ETM+ data. *International Journal of Wildland Fire*, 17, 527. doi:10.1071/WF08038
- Vosselman, G. (2000). Slope based filtering of laser altimetry data. *International Archives of Photogrammetry and Remote Sensing*, 33, 935–942.
- Vosselman, G., & Maas, H.-G. (2010). *Airborne and Terrestrial Laser Scanning*. Whittles Publishing, Dunbeath, Escocia, 318 pp:
- Wang, C., Qi, J., & Cochrane, M. (2005). Assessment of Tropical forest degradation with Canopy Fractional Cover from Landsat ETM+ and IKONOS imagery. *Earth Interactions*, 9, 1–18. doi:10.1175/EI133.1
- Wang, H., Wang, S., Chen, Q., Jin, W., & Sun, M. (2015). An improved filter of progressive TIN densification for LiDAR point cloud data. *Wuhan University Journal of Natural Sciences*, 20, 362–368. doi:10.1007/s11859-015-1106-9
- Watt, M.S., Adams, T., Watt, P., & Marshall, H. (2013). Influence of stand and site conditions on the quality of digital elevation models underlying New Zealand forests. *New Zealand Journal of Forestry Science*, 43, 1–9. doi:10.1186/1179-5395-43-5
- Wehr, A., & Lohr, U. (1999). Airborne laser scanning—an introduction and overview. *ISPRS Journal of Photogrammetry and Remote Sensing*, 54, 68–82. doi:10.1016/s0924-2716(99)00011-8
- Wulder, M.A., Bater, C.W., Coops, N.C., Hilker, T., & White, J.C. (2008). The role of LiDAR in

- sustainable forest management. *The Forestry Chronicle*, 84, 807–826. doi:10.5558/tfc84807-6
- Wulder, M.A., White, J.C., Alvarez, F., Han, T., Rogan, J., & Hawkes, B. (2009). Characterizing boreal forest wildfire with multi-temporal Landsat and LiDAR data. *Remote Sensing of Environment*, 113, 1540–1555. doi:10.1016/j.rse.2009.03.004
- Wulder, M.A., White, J.C., Nelson, R.F., Næsset, E., Ørka, H.O., Coops, N.C., Hilker, T., Bater, C.W., & Gobakken, T. (2012). LiDAR sampling for large-area forest characterization: A review. *Remote Sensing of Environment*, 121, 196–209. doi:10.1016/j.rse.2012.02.001
- Yu, X., Hyyppä, J., Holopainen, M., & Vastaranta, M. (2010). Comparison of Area-Based and Individual Tree-Based methods for predicting plot-level forest attributes. *Remote Sensing*, 2, 1481–1495. doi:10.3390/rs2061481
- Yu, X., Hyyppä, J., Vastaranta, M., Holopainen, M., & Viitala, R. (2011). Predicting individual tree attributes from airborne laser point clouds based on the random forests technique. *ISPRS Journal of Photogrammetry and Remote Sensing*, 66, 28–37. doi:10.1016/j.isprsjprs.2010.08.003
- Zhang, K., Chen, S., Whitman, D., Shyu, M., Yan, J., Zhang, C., & Member, S. (2003). A progressive morphological filter for removing nonground measurements from Airborne LiDAR Data. *IEEE Transactions on Geoscience and Remote Sensing*, 41, 872–882. doi:10.1109/TGRS.2003.810682
- Zhang, K., & Whitman, D. (2005). Comparison of three algorithms for filtering airborne lidar data. *Photogrammetric Engineering and Remote Sensing*, 71, 313–324. doi:10.14358/PERS.71.3.313
- Zhang, J., & Lin, X. (2013). Filtering airborne LiDAR data by embedding smoothness-constrained segmentation in progressive TIN densification. *ISPRS Journal of Photogrammetry and Remote Sensing*, 81, 44–59. doi:10.1016/j.isprsjprs.2013.04.001
- Zhao, X., Guo, Q., Su, Y., & Xue, B. (2016). Improved progressive TIN densification filtering algorithm for airborne LiDAR data in forested areas. *ISPRS Journal of Photogrammetry and Remote Sensing*, 117, 79–91. doi:10.1016/j.isprsjprs.2016.03.016

APÉNDICE

En el siguiente cuadro resumen se especifican las referencias de los artículos que constituyen el cuerpo de la tesis, así como el factor de impacto (FI, *Journal Citation Reports* 2015) de las revistas, las áreas temáticas correspondientes a las publicaciones y la justificación de la contribución del doctorando al tratarse de trabajos realizados en coautoría. Estos trabajos de investigación no formarán parte de ninguna otra tesis en modalidad de compendio de publicaciones.

Publicación y justificación de la contribución del doctorando

Montealegre, A.L., Lamelas, M.T., de la Riva, J. 2015. Comparison of open source LiDAR filtering algorithms in a Mediterranean forest environment. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 8: 4072-4085. doi:10.1109/JSTARS.2015.2436974.

Antonio Luis Montealegre ha realizado el grueso del trabajo, desarrollando la mayor parte del análisis estadístico y espacial, siendo el responsable último de la redacción de los contenidos, discusión y conclusiones. La Dra. María Teresa Lamelas figura como coautora en calidad de responsable del desarrollo metodológico y del análisis estadístico de la investigación; asimismo ha colaborado en la revisión de los resultados y en el trabajo de campo junto con el doctorando. El Dr. Juan de la Riva ha colaborado en la revisión de los contenidos.

FI	FI 5 años	Categoría de JCR®	Número orden	Cuartil
2,145	2,613	Engineering, electrical & electronic	61 de 257	Q1
		Remote sensing	11 de 28	Q2
		Imaging science & photographic technology	7 de 24	Q2
		Geography, physical	21 de 49	Q2

Publicación y justificación de la contribución del doctorando

Montealegre, A.L., Lamelas, M.T., de la Riva, J. 2015. Interpolation routines assessment in ALS-derived Digital Elevation Models for forestry applications. *Remote Sensing*, 7: 8631-8654. doi:10.3390/rs70708631.

Antonio Luis Montealegre ha realizado la mayor parte de la investigación, desarrollando el grueso del análisis estadístico y espacial, siendo el responsable último de la redacción de los contenidos, discusión y conclusiones. La Dra. María Teresa Lamelas ha tenido la idea original de la investigación y figura como coautora en calidad de responsable del desarrollo metodológico y del análisis estadístico de la investigación; asimismo ha colaborado en la revisión de los resultados y en el trabajo de campo junto con el doctorando. El Dr. Juan de la Riva figura como coautor habiendo también colaborado en la revisión de los contenidos.

FI	FI 5 años	Categoría de JCR®	Número orden	Cuartil
3,036	3,278	Remote sensing	5 de 28	Q1

Publicación y justificación de la contribución del doctorando

Montealegre, A.L., Lamelas, M.T., de la Riva, J., García-Martín, A., Escribano, F. 2016. Use of low point density ALS data to estimate stand-level structural variables in Mediterranean Aleppo pine forest. *Forestry*, 89 (4): 373-382. doi:10.1093/forestry/cpw008.

Antonio Luis Montealegre ha realizado el grueso de trabajo, desarrollando la mayor parte del análisis estadístico y espacial, siendo el responsable último de la redacción de los contenidos, discusión y conclusiones. La Dra. María Teresa Lamelas ha sido la responsable del desarrollo metodológico y del análisis estadístico de la investigación. Todos los coautores han participado de las tareas de trabajo de campo y han colaborado en la revisión de los contenidos.

FI	FI 5 años	Categoría de JCR®	Número orden	Cuartil
1,921	2,184	Forestry	13 de 66	Q1

Publicación y justificación de la contribución del doctorando

Montealegre, A.L., Lamelas, M.T. Tanase, M., de la Riva, J. 2014. Forest fire severity assessment using ALS data in a Mediterranean environment. *Remote Sensing* 6: 4240-4265. doi:10.3390/rs6054240.

Antonio Luis Montealegre ha realizado el grueso del trabajo, desarrollando la mayor parte del análisis estadístico y espacial, siendo el responsable último de la redacción de los contenidos, discusión y conclusiones. La Dra. María Teresa Lamelas figura como coautora en calidad de responsable del desarrollo metodológico y del análisis estadístico de la investigación. El Dr. Mihai Tanase ha proporcionado los datos de campo y el Dr. Juan de la Riva ha tenido la idea original del trabajo. Todos los coautores han colaborado en la revisión de los resultados.

FI	FI 5 años	Categoría de JCR®	Número orden	Cuartil
3,036	3,278	Remote sensing	5 de 28	Q1

Por otro lado, se han incluido en la tesis dos comunicaciones a congresos que no cuentan con factor de impacto. Al tratarse de trabajos realizados en coautoría, a continuación se detalla la contribución del doctorando.

Publicación y justificación de la contribución del doctorando

Montealegre, A.L., Lamelas, M.T., de la Riva, J., García-Martín, A., Escribano, F. 2015. Assessment of biomass and carbon content in a Mediterranean Aleppo pine forest using ALS data. In *Proceedings of the 1st Int. Electron. Conf. Remote Sens.*, 22 June–5 July 2015; Sciforum Electronic Conference Series, Vol. 1, d004. doi:10.3390/ecrs-1-d004.

Antonio Luis Montealegre ha realizado el grueso del trabajo, desarrollando la mayor parte del análisis estadístico y espacial, siendo el responsable último de la redacción de los contenidos, discusión y conclusiones. La Dra. María Teresa Lamelas ha sido la responsable del desarrollo metodológico y del análisis estadístico de la investigación. Todos los coautores han participado de las tareas de trabajo de campo y han colaborado en la revisión de los contenidos.

Sistema de revisión de comunicación por expertos doble ciego. Presidente del comité científico: Ioannis Gitas – Universidad Aristóteles de Salónica (Grecia).

Publicación y justificación de la contribución del doctorando

Montealegre, A.L., Lamelas, M.T., de la Riva, J., García-Martín, A., Escribano, F. 2015. Cartografía de modelos de combustible mediante combinación de imágenes LiDAR, SAR y ópticas en el Centro de adiestramiento “San Gregorio”. In Asorey Cacheda, R., Devesa Rey, R., Solla Carracelas, M.M., Pousada Carballo, J.M. (Eds), *Congreso Nacional de i+d en Defensa y Seguridad, DESEi+d 2015*. Centro Universitario de la Defensa de Marín, Marín, pp. 1185-1192 (ISBN:978-84-944537-0-0).

Antonio Luis Montealegre ha realizado el grueso del trabajo, desarrollando la mayor parte del análisis estadístico y espacial, siendo el responsable último de la redacción de los contenidos, discusión y conclusiones. La Dra. María Teresa Lamelas y el Dr. Alberto García-Martín han sido responsables del diseño metodológico. Todos los coautores han participado de las tareas de trabajo de campo y han colaborado en la revisión de los contenidos.

Sistema de revisión de comunicación por expertos doble ciego. Presidenta del comité científico: María Mercedes Solla Carracelas – Centro Universitario de la Defensa de Marín (España).